Utilizing YOLOv6 Deep Learning Algorithm Trained on Custom Dataset to Accurately Count Sugar Beet Plants within Agricultural Fields

P. Friedrich[1] & G. Schimmer[1]

[1] *M.Sc. Geography Student at Heidelberg University*

Contact

philipp.friedrich@stud.uni-heidelberg.de

gustav.schimmer@stud.uni-heidelberg.de

Keywords: *Sugar Beet, Deep Learning, YOLOv6, Object Detection, Agriculture, Crop Science, Crop Monitoring, Photogrammetry*

Abstract

Sugar beet cultivation plays a significant role in modern agricultural practices, necessitating precise management strategies to optimize yield and profitability. Central to this optimization is the accurate counting of sugar beet plants within cultivated fields directly after emergence, a task traditionally performed through labor-intensive manual methods. However, advancements in technology, particularly in the realm of deep learning and image analysis, offer promising alternatives to streamline this process.

This paper presents a novel approach leveraging deep learning techniques, specifically the YOLOv6 algorithm, to automate the detection and counting of sugar beet plants in agricultural fields using drone or smartphone imagery. By harnessing the capabilities of computer vision, our method aims to overcome the limitations of conventional counting methods, providing farmers with a more efficient and accurate means of assessing plant populations.

The project shows the potential of YOLOv6 algorithm for crop detection and addresses the influence of different image resolution on model performance.

1. Introduction

Sugar beet cultivation plays a significant role in modern agricultural practices in several regions in Europe, necessitating precise management strategies to optimize yield and profitability. Central to this optimization is the accurate counting of sugar beet plants within cultivated fields directly after emergence, a task traditionally performed through labor-intensive manual methods. However, advancements in technology, particularly in the realm of deep learning and image analysis, offer promising alternatives to streamline this process (Kitano et al. 2019, Gao et al. 2020, Zhang et al. 2020, Divyanth et al. 2023).

This paper presents a novel approach leveraging deep learning techniques, specifically the YOLOv6 algorithm, to automate the detection and counting of sugar beet plants in agricultural fields using drone or smartphone imagery. By harnessing the capabilities of computer vision, our method aims to overcome the limitations of conventional counting methods, providing farmers with a more efficient and accurate means of assessing plant populations.

The conventional method of plant counting involves sampling a small portion of the field and extrapolating the results - a process prone to inaccuracies and labor inefficiencies. In contrast, our approach enables comprehensive plant counting across the entire field, facilitating informed decision-making regarding replanting schedules and pesticide dosing (Kitano et al. 2019).

This project aims to enhance crop yields, minimize pesticide usage, and promote sustainable agricultural practices through precision farming techniques. As a result, it can also support the achievement of the United Nations' Sustainable Development Goal 2: Zero Hunger.

Key objectives of this study include:

- Developing a robust methodology for automated sugar beet plant detection and counting using YOLOv6.
- Assessing the feasibility and accuracy of employing drones or smartphone imagery for field-level plant monitoring.
- Exploring the potential economic and agronomic benefits of implementing our automated counting approach in sugar beet cultivation.
- Through this research, we aim to contribute to the advancement of precision agriculture practices, enhancing the sustainability and productivity of sugar beet farming operations.

2. State of the Art

In recent years, there has been a surge in research focused on developing advanced methodologies for plant segmentation using deep learning techniques. These methodologies aim to address the challenges associated with traditional plant monitoring methods

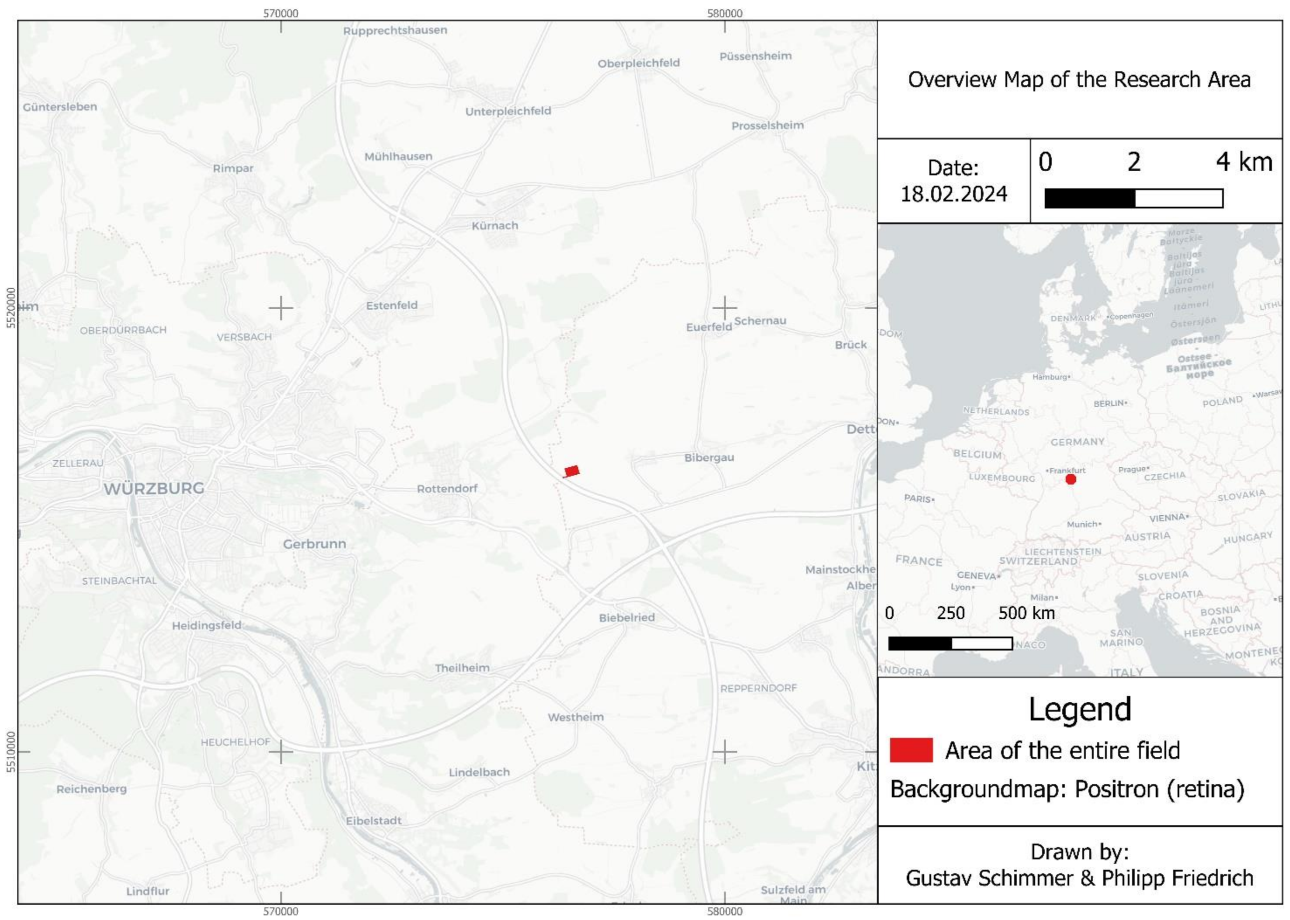

Figure 1: Location of the research area near the city of Würzburg, Germany (own figure).

and offer novel approaches for accurate and efficient plant analysis.

Gao et al. (2020) reviewed the latest advancements in deep learning techniques for crop stress diagnosis. The study compiled various sensor tools and deep learning principles relevant to plant stress phenotyping, including classification, object detection, and segmentation.

Kitano et al. (2019) developed techniques for corn plant counting and automation using UAV images and deep learning algorithms. The study utilized low-cost UAV platforms equipped with RGB sensors to capture images of corn crops. The researchers were able to accurately count corn plants in large-scale production fields, overcoming the limitations of labor-intensive manual counting methods.

In contrast deep learning methods can also be used for plant segmentation in diverse environments. Zhang et al. (2020) proposed a novel deep learning method, the Scale Sequence Residual U-Net (SS Res U-Net), for identifying and mapping individual plants in high-elevation ecosystems using UAV imagery. By applying the SS Res U-Net to identify and map dominant plant species in paramos ecosystems, high classification accuracy and robustness could be achived.

Such studies highlight the significant advancements in plant segmentation methodologies enabled by deep learning technology. By using these approaches, researchers can overcome traditional challenges in plant monitoring and contribute to more efficient and sustainable agricultural practices. Future research in this field is expected to focus on further refining deep learning algorithms, enhancing data acquisition techniques, and expanding the application of plant segmentation methods to various ecosystems and crop types.

3. Data & Methods

3.1. Research Area and Data Acquisition

The images utilized in this study were captured on an agricultural field located near the city of Würzburg in Bavaria. The field was cultivated with sugar beets during the year 2023, and the sowing process occurred on April 13, 2023. Throughout the entire field, covering an expansive area of 6.4 hectares, images were obtained from two distinct study areas, each spanning approximately 9 square meters and 5 rows of crops. Image acquisition occurred at various time steps throughout the growing period. Figure 1 provides an overview of the specific location.
The entire area surrounding the site is intensively cultivated, with winter wheat being the predominant crop. Additionally, corn, sugar beets, canola, and spelt are also grown in the region. The average annual precipitation amount for the climate normal period from 1981 to 2010, as recorded at the Würzburg climate station, is 596.1 millimeters. The peak of precipitation occurs in July, while the minimum is observed in September (DWD, 2022).

The images were manually captured using a Huawei P30 Lite smartphone and a standard selfie stick. The smartphone was positioned at an approximate height of 100 centimeters, resulting in an image footprint of 2x1 meter and a ground sampling distance (GSD) of 0.5 millimeters. For each time step, up to 120 images were taken at each study site, with an overlap ranging from 50% to 75%. For the model training, only images captured on May 14, 2023, were used, as the algorithm is primarily intended to recognize and count sugar beet plants in the earliest stages. The mentioned date was the earliest possible with available images. At this time the individual plants were predominantly in the four-leaf stage.

3.2. Methods

The general workflow of the project is shown in Figure 2 and will be explained in the following. To train a deep learning model, a substantial dataset of labeled images is essential. We therefore used the images acquired by smartphone as described in the chapter before. Preprocessing involved two major steps. On the one side an orthophoto was generated using Agisoft Metashape software und measured ground control points (GCPs). The second step included the preparation of a custom dataset for YOLOv6 algorithm training. In this instance, labels for individual sugar beet plants were required. These image labels were manually generated using the open-source and freely available web application

Alpha Make Sense (makesense.ai). This tool provides output file formats compatible with the YOLOv6 model. Both the images and labels are utilized in the subsequent model training process.

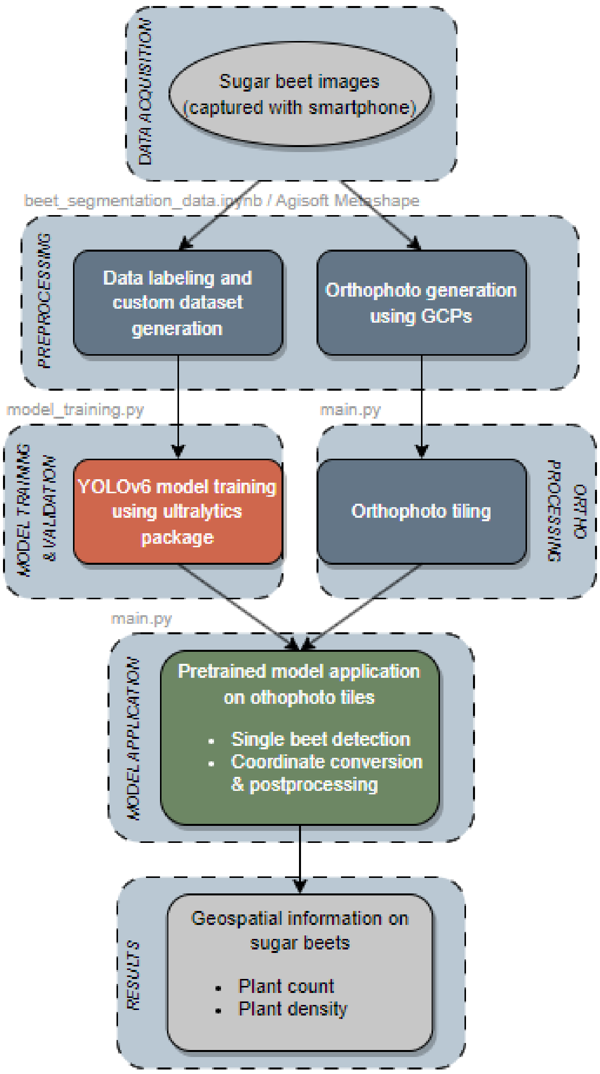

Figure 2: General Workflow of the project (own figure).

To expedite computational time, streamline processing steps, and enhance overall efficiency, we tried to reduce the original image resolution of 4000x3000 pixels. Consequently, we resampled the images to explore whether employing lower resolutions significantly impacted the model's accuracy. Ultimately, we opted for the intermediate resolution among those examined, as it appeared to strike the best compromise between accuracy and processing speed. The final resolution chosen was 2000x1500 pixels per image (Figure 3). To achieve this resolution, we utilized the "resize" function from the `cv2` package.

YOLOv6 Model training was done within Python programming language using `ultralytics` package. The dataset comprises 1510 images, each with a pixel size of 512x512, adhering to the requirement of YOLOv6 for image sizes to be multiples of 32. To ensure robust model performance, we partitioned the dataset into training and validation sets with an 80/20 split ratio. Training was performed over 400 epochs, with the best results achieved at epoch 194.

The implementation and application of the model were also executed in Python, using various packages including geospatial libraries like `GDAL`, `rasterio` or `shapely`. Given the YOLOv6 model's requirement for images of consistent dimensions, we employed a tiling approach on the orthophoto dataset while preserving the geospatial information of each tile. Following the model's integration and subsequent detection of sugar beets across orthophoto tiles, we extracted the bounding boxes of each image. These bounding boxes were then translated from pixel coordinates to geographical coordinates, ensuring alignment with real-world spatial context. To ensure accuracy and only have one bounding box per plant, a post-processing step was conducted. Here, individual bounding boxes were buffered with a radius of 2.5 cm, merged and rebuffered to their original dimensions. Finally, the resulting bounding boxes within the defined research area, along with their centroids, were exported for further analysis and interpretation.

Figure 3: Comparison of different image resolutions (own figure).

4. Results

Model Training

The model was trained using `ultralytics` package in Python. 400 epochs were applied for model training. The best result was achieved after 194 epochs. The number of True Positives (TP) generated from the model is 385, in contrast to 54 False Positives (FP) and 26 False Negatives (FN) (Figure 4). The Precission after 194 epochs of training is 0.87, the Recall is 0.89 and the mAP50 is 0.92, the mAP50-95 is 0.51.

Application on Research Area

For model application, we used the previously trained model to identify sugar beets within an orthophoto and a specified research area of approximately 9 m². The orthophoto was tiled into 550 individual images. The results are shown in Figure 5. Within the orthophoto, there were 69 sugar beet plants, out of which the model accurately identified 66. Three sugar beet plants went undetected, and tree weed plants were falsely identified as sugar beet.

5. Discussion

In this project, we trained a YOLOv6 algorithm on a custom dataset to detect individual sugar beet plants in agricultural fields. The results demonstrate a robust model performance, including some minor limitation. The confidence levels of the model predictions range from 0.5 to 0.9 with the majority hovering around 0.7. This suggests possible improvements in model accuracy using a larger and more diverse training dataset. However, the results obtained from the orthophoto of our research area are quite solid. Which is possibly related to orthophoto processing using the same images as were used for model training. Throughout the model development

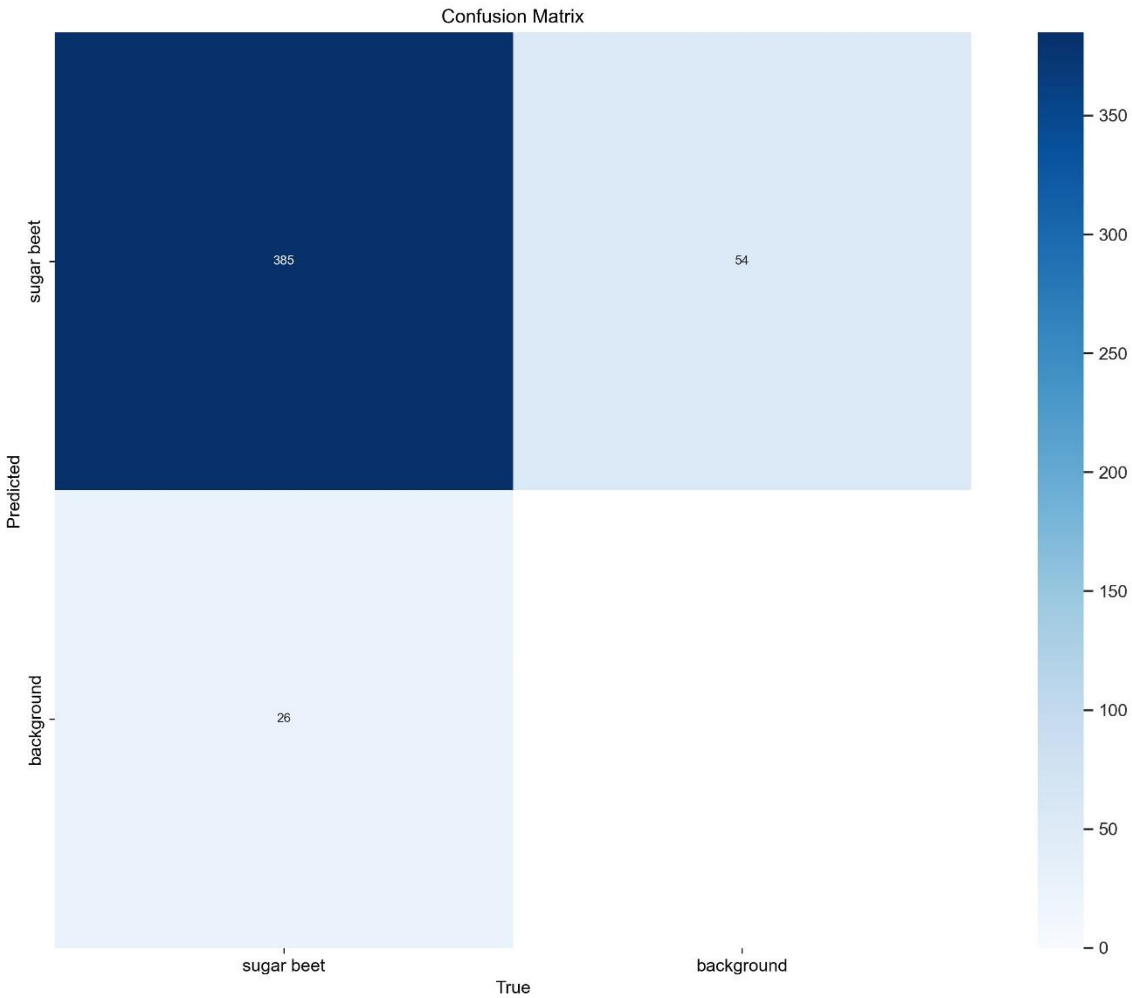

Figure 4: Confusion matrix of the YOLOv6 deep learning model trained with 194 epochs (own figure).

process, we gained several insights. The key observations will be briefly discussed.

Image Resolution Optimization

In our methodology, we explored the impact of image resolution on the performance of the YOLOv6 model for sugar beet plant detection. Through resolution testing, we aimed to strike a balance between computational efficiency and model accuracy. Our findings indicate that while reducing image resolution can expedite processing time, excessively low resolutions compromise detection accuracy. The selected resolution of 2000x1500 pixels per image demonstrated optimal performance, enabling efficient processing without significant loss in accuracy. However, it's worth noting that further experimentation may be warranted to evaluate performance under varying field conditions and vegetation densities.

Weed Detection and Model Limitations

An important consideration in automated plant detection is the potential interference from weeds, which may confound the model's ability to accurately identify sugar beet plants. While our YOLOv6 model successfully detects and counts sugar beet plants, distinguishing between sugar beet and weed species remains a challenge. This limitation underscores the need for ongoing refinement and augmentation of the model to improve its capacity for discriminating between target crops and undesired vegetation. Future iterations of the model could incorporate additional training data encompassing diverse weed species to enhance its robustness under real-world field conditions.

Overfitting and Generalization

An inherent challenge in model training is the risk of overfitting to the training dataset, thereby

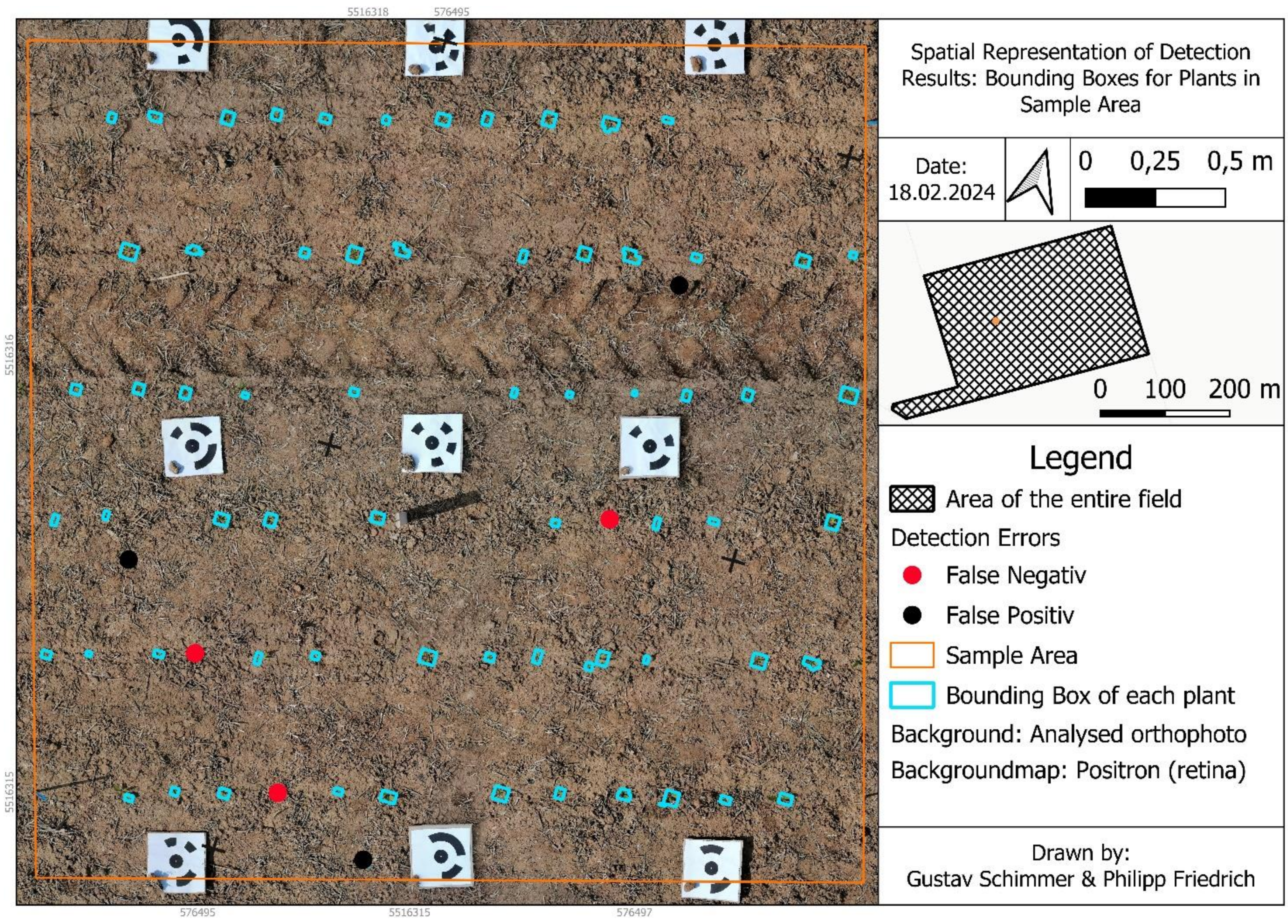

Figure 5: Spatial representation of detection results in research area (own figure).

compromising the model's ability to generalize to unseen data. While our study focused on a specific geographic location and growing season, the risk of overfitting necessitates cautious interpretation of results. Cross-validation techniques and external validation using independent datasets could provide valuable insights into the generalizability and robustness of the model across different environmental conditions and cultivation practices. Therefore, there is a strong need in testing the model in other geographical regions.

Practical Applicability and Future Directions

Despite its simplicity, our approach presents a viable solution for automating plant counting in sugar beet cultivation. By leveraging readily available technology such as smartphones and drones, farmers can streamline field monitoring processes, optimize resource allocation, and make informed management decisions. However, there exist avenues for further enhancement, including the integration of multispectral or hyperspectral imaging for advanced crop analysis and the development of sophisticated algorithms for precise weed identification and management. The model should also be trained to detect sugar beet plants in earlier stages of growth. Providing information on plant emergence and density to farmers at an early stage is crucial, as it allows for greater flexibility and informed decision-making. Moreover, expanding the scope of the study to encompass diverse geographical regions and crop types would facilitate broader applicability and validation of the proposed methodology.

6. Conclusion

In this study, we have demonstrated the potential of leveraging deep learning techniques, specifically the YOLOv6 algorithm, for automating the detection and counting of sugar beet plants in agricultural fields using drone or smartphone imagery. Our approach represents a significant advancement over traditional manual counting methods, offering farmers a more efficient and accurate means of assessing plant populations.

Through careful experimentation and analysis, we optimized image resolution to strike a balance between computational efficiency and model accuracy, ultimately selecting a resolution of 2000x1500 pixels per image. This resolution enabled efficient processing without compromising detection accuracy, paving the way for practical implementation in real-world agricultural settings.

While our model shows promising results, it is not without limitations. Weed interference poses a significant challenge to accurate plant detection, highlighting the need for ongoing refinement and augmentation of the model. Future research directions may include the integration of multispectral or hyperspectral imaging for advanced crop analysis, as well as the development of sophisticated algorithms for precise weed identification and management.

Despite these challenges, our study represents a crucial step towards the adoption of precision agriculture practices in sugar beet cultivation. By embracing technological innovation and interdisciplinary collaboration, we can enhance the sustainability and productivity of crop production systems, ultimately contributing to global food security and agricultural resilience. This also relates to United Nations Sustainable Development Goals. Therefore, this study can also be seen as a contribution to SDG 2: Zero Hunger.

In conclusion, our findings underscore the transformative potential of deep learning in revolutionizing agricultural practices, and we look forward to further advancements in this exciting field of research.

References

Divyanth, L. G., Ahmad, A., & Saraswat, D. (2023). A two-stage deep-learning based segmentation model for crop disease quantification based on corn field imagery. *Smart Agricultural Technology*, 3, 100108.

Gao, Z., Luo, Z., Zhang, W., Lv, Z., & Xu, Y. (2020). Deep learning application in plant stress imaging: a review. *AgriEngineering*, 2(3), 29.

Kitano, B. T., Mendes, C. C., Geus, A. R., Oliveira, H. C., & Souza, J. R. (2019). Corn plant counting using deep learning and UAV images. *IEEE Geoscience and Remote Sensing Letters*.

Zhang, C., Atkinson, P. M., George, C., Wen, Z., Diazgranados, M., & Gerard, F. (2020). Identifying and mapping individual plants in a highly diverse high-elevation ecosystem using UAV imagery and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, 280-291.