Plant Counts Using Computer Vision And Deep Learning

Ameer Tamoor Khan
University of Copenhagen
atk@plen.ku.dk

December 7, 2023

Abstract

This study explores the application of advanced computer vision techniques in straw estimation by counting the number of sepals of spring wheat and barley using Robotti ground images and aerial drone imagery. We utilized high-resolution Robotti images for initial model training due to their detailed visual data and ease of annotation. The study began with classical image processing methods, including HSV and LAB colormaps, for sepal count but quickly transitioned to deep learning due to the complexity of the data. Our approach centered on a supervised learning model, where a dataset of 50 images was annotated for sepal counting among other features using a tool "Hasty." We chose Mask R-CNN for its excellence in instance segmentation, training it on this annotated dataset. The model demonstrated exceptional accuracy and adaptability in various test conditions. Additionally, we developed a user-friendly web interface to facilitate interaction with our models, allowing for model testing, training, and custom analysis. Currently, it runs locally, however, upon completion it will available for use. Our future direction includes expanding the dataset, refining model precision, and broadening the scope to include additional crop types and exploring other advanced models.

Dataset Overview

Our study primarily utilized high-resolution Robotti images, focusing on spring wheat and barley. These images, taken close to the ground, offer detailed visual data that is crucial for our analysis. The proximity to the crops allows for easier and more precise annotation, making these Robotti images an ideal building block for our project. It will simplify the initial training phase of our model.

Robotti images are relatively easier to analyze and annotate compared to drone images, which often present complexity due to their higher vantage point and the vastness of the area they capture. By annotating a large number of Robotti images, we can do so more conveniently and efficiently, setting a solid foundation for the model. Once the model has been effectively trained on Robotti images, we will incrementally introduce drone images. This step involves annotating a select number of drone images to fine-tune the model, ensuring that it adapts and performs well with the more complex aerial data. This phased approach allows for gradual model improvement and ensures robust performance across different types of imagery.

Classical Methodologies

Initially, we employed classical image processing techniques, including HSV and LAB colormaps, to estimate straw. These methods were aimed at isolating and identifying key crop features. However, the complexity of the dataset posed significant challenges, rendering these techniques insufficient for achieving our objective of accurate sepal counting.

Adoption of Deep Learning

To overcome the limitations of classical methods, we shifted our focus to deep learning, particularly image segmentation. Our goal was to segment wheat and barley sepals accurately from the images. We adopted a supervised learning approach, manually annotating a set of 50 images – 18 from wheat and 32 from barley. The annotation process involved categorizing five classes: sepals, weed thistle, weed monocot, weed dicot, and grass. This detailed categorization was crucial for training the model to differentiate sepal features effectively.

We used "Hasty" tool that assists with the annotation of images for training computer vision models. Utilizing Hasty's intuitive AI-powered platform, we aimed to accurately segment sepals from wheat and barley in our dataset. This tool facilitated a more efficient and precise annotation process, thereby enabling our model to effectively distinguish sepal features and vastly improving our supervised learning approach.

- **Initial Dataset Composition:**
  - Total images: 50
    - Wheat images: 12
    - Barley images: 38
- **Class Categories for Annotation:**
  - Sepal
  - Weed thistle
  - Weed monocot
  - Weed dicot
  - Grass
- **Dataset Allocation:**
  - Training dataset: 45 images
  - Testing dataset: 5 images
    - Wheat: 3
    - Barley: 2

Model Training and Implementation

For the deep learning model, we chose Mask R-CNN, an advanced model that excels in instance segmentation, capable of identifying and outlining each object in an image. It extends the Faster R-CNN framework by adding a branch for segmentation, allowing it to predict object boundaries with high precision. Our model was trained on annotated images, fine-tuning it to accurately distinguish plant parts such as sepals from their surroundings and from weeds.

Results

Our implementation of Mask RCNN yielded exceptional results. We achieved a high accuracy rate of 98.71% and a significantly low loss of 0.1800, surpassing traditional estimation methods. The mean Average Precision (mAP) for both segmentation and bounding box demonstrated the model's efficacy in capturing the nuances of wheat and barley crops. Moreover, our tests across various drone image heights (2m, 4m, 8m) confirmed the model's robustness and adaptability in different environmental conditions.

Table 1 shows the detailed accuracy of sepal counting for wheat and barley. Figure 1 provides visual representation of model testing on wheat and barley images.

Type	**Ground Truth**	**Predicted**	**Error %**
Wheat	122	129	5.00%
Barley	317	319	0.60%
Wheat	309	276	10.0%
Wheat	125	128	2.00%
Barley	183	204	11.0%

Table 1: Model Testing On Five Test Images.

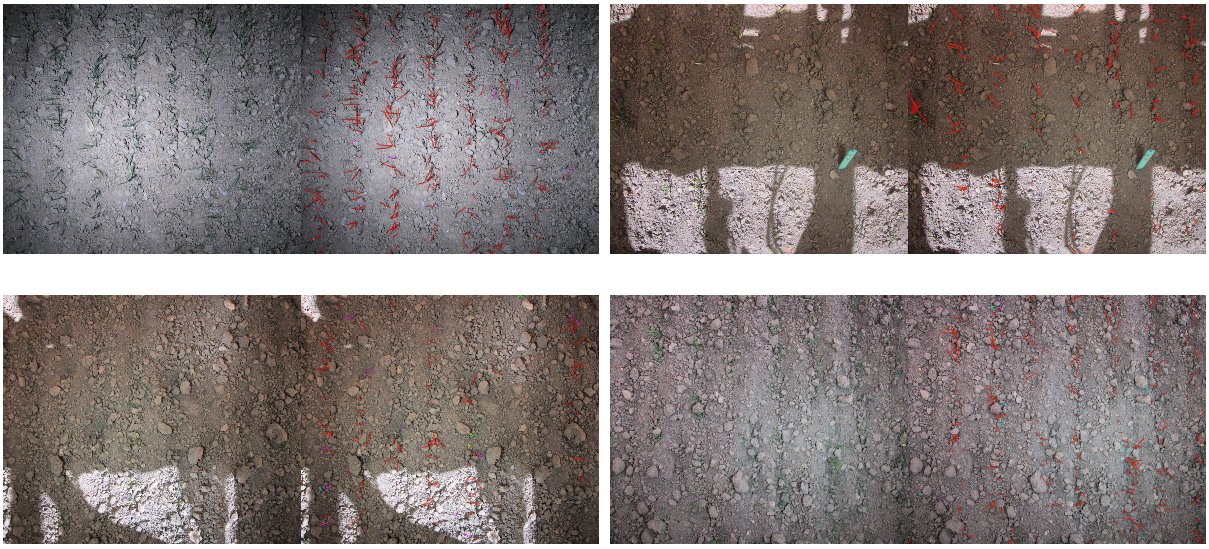

Figure 1: The Sample Test Images effectively demonstrate the model's proficiency in distinguishing sepals from the surrounding background and weeds.

Web Interface for Model Interaction

To facilitate user interaction with our deep learning models, we have developed a comprehensive web interface. This platform is designed to be intuitive and user-friendly, enabling breeders and researchers to directly engage with the model's capabilities. Currently, the web-inerfdace is running on a local machine. At present, the web interface is operational on a local machine. Upon completion and readiness for release, it will be made available to breeders for testing purposes. The interface is comprised of three core components:

- **Model Testing**: This feature allows users to upload images for sepal counting. The model processes the uploaded images and provides a count of various elements, including the number of sepals, weed thistle, monocut, dicot, and grass. This instant analysis aids in assessing the model's precision and real-world applicability.
- **Model Training**: The interface also serves as a gateway for model refinement. By allowing the upload of new image datasets, users can contribute to the training process, enhancing the model's learning and improving its subsequent predictions.
- **Custom Tester**: A specialized tool within the interface enables breeders to calculate sepals in a selected part of the image. This targeted approach allows for detailed analysis, giving breeders the flexibility to focus on specific areas of interest within the crop imagery.

This web interface represents a significant step towards democratizing the use of advanced computer vision in agriculture, making powerful analytical tools accessible to experts in the field.

Future Directions

As we look to the future, our project is set to embrace a series of strategic enhancements aimed at refining our deep learning models and expanding their scope of application. These advancements are planned with the intention of not only improving the accuracy of our current models but also extending their utility across a broader range of crop types and developmental stages. Our forward-looking strategy includes the following key initiatives:

- Expand our dataset by annotating more images to improve the model's accuracy and robustness.
- Focus on individual training for wheat and barley to refine model precision and tailor it to specific crop characteristics.
- Include additional crop types such as winter wheat, winter triticale, and winter rye to broaden the model's scope and adaptability.
- Explore the annotation of drone images to enhance training datasets, capturing a wider array of field conditions and plant health indicators.
- Experiment with other promising models in computer vision and deep learning to continuously improve performance and stay abreast of technological advancements.

Conclusion

Our research successfully demonstrated the efficacy of applying deep learning techniques, particularly Mask R-CNN, in agricultural image analysis. The transition from classical image processing to deep learning was pivotal in overcoming the challenges posed by complex datasets. Our model showed high accuracy in sepal counting for both wheat and barley, outperforming traditional methods. The implementation of a phased approach, beginning with Robotti images and gradually integrating drone imagery, proved effective in enhancing the model's robustness. The development of a comprehensive web interface marks a significant step in democratizing advanced computer vision tools for agricultural purposes. Looking ahead, our focus will be on dataset expansion, individual crop-specific model refinement, inclusion of additional crop types, and continuous exploration of emerging technologies in computer vision. This research lays the groundwork for future innovations in precision agriculture, aiming to improve crop analysis and management through advanced computational methods.