SeeTree Optimizes Disease Detection with AI Data Management

POWERED BY DATALOOP

About SeeTree

SeeTree is a unique company that provides visibility into the health records and productivity of any individual tree at any time, and overtime. The result is a transformation of the way trees are managed, leading to a significant increase in profitability. Their efforts revolutionized farming for trees, taking the tree world and digitizing it, allowing the farmer to understand every tree he has and optimize his job with tools to better maximize the outcomes.

SeeTree provides growers with the data needed on a macro and micro level for optimized farming. Their existing customer base has expressed that SeeTree answers their longtime need for true precision farming.

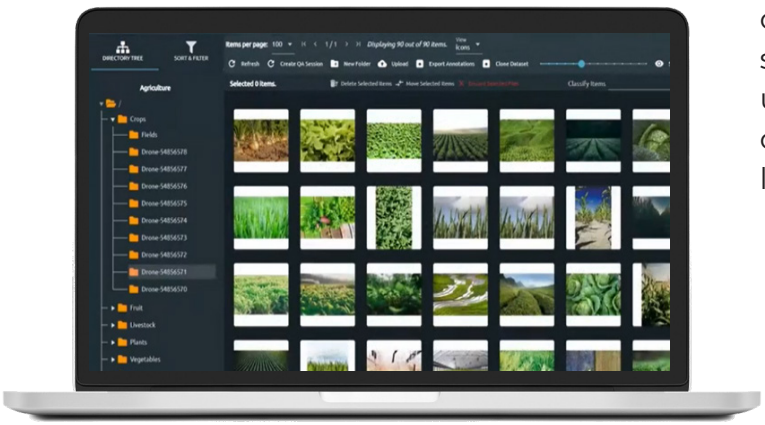
The Farmers Ongoing Challenge To Have “All Eyes” on Their Fields

The past few years have given rise to farmers incorporating computer vision models in order to improve farming applications ranging from monitoring crops to producing livestock and aquaculture. Farmers need to be able to detect problems as early as possible, if they want to succeed. Up until today, they’ve used their eyes and their feet to keep track of their crops.

Today, they still have to rely on their eyes and feet to see what is in their groves, but threats are now able to be picked up much sooner than they would if the work was performed manually. Which in turn speeds up the process and increases their quality as well as productivity. For example, bacteria has the potential to wipe out an entire farm, and replacing one single tree takes 5 years to grow! Therefore, it’s essential for a farmer to have a comprehensive service of combining boots on the ground, and high-resolution multi-dimensional sensing imagery utilizing drones, sensors, and special vehicles to collect precise data.

A Data-Driven Solution for the Farmer

We’ve established that one of the prominent challenges a farmer faces is detecting diseases in their crops. The ideal situation is for a farmer to be able to detect crop diseases as quickly as possible in order to reduce the spread of these crop diseases. In some cases, if a single tree is found to be infected, that tree is removed and cut at its roots in order to halt the spread to the surrounding trees. If you’re talking about 100 trees, it is possible to inspect them from the ground, however, when you reach 2 million trees, the only way you can monitor them effectively is with AI.



AI is the data-driven solution for the farmer’s dilemma, giving them automating “eyes” in the field. This increases the farmer's ability to control and manage on a higher level without compromising on quality. However, the problem with AI is that it also requires human eyes. In order to generate AI, you need humans in the annotation process. So we’ve really just taken the eyes from the field and moved them to the “laptop”. AI allows you to scale the process, and this is where automation becomes crucial.

SeeTree mainly uses Dataloop to annotate their trees. In particular, they're focusing on bind weeds that climb on trees to get sunlight. SeeTree is looking to identify under-performing trees, locate strong clusters and potentially weed out any lingering tree diseases.

SeeTree Speeds Up Data Workflows with Dataloop's SDK and Data Query Language

At SeeTree, they're collecting, scanning, and analyzing hundreds of millions of trees in order to provide the most comprehensive and dependable metrics per tree. Until recently, SeeTree managed annotations in a very manual manner, but now with Dataloop's developer-friendly unique SDK they're able to accelerate their entire data preparation flow from end to end. They're able to ingest data seamlessly, can filter, sort, and query in order to pick which trees are appearing in specific images from their data sets.

How does this work?

STEP #1
Sample data is seamlessly ingested into Dataloop

STEP #2
Labeling task is generated and sent to human validation

STEP #3
Human in the loop validation of infestation symptoms are used for retraining the model

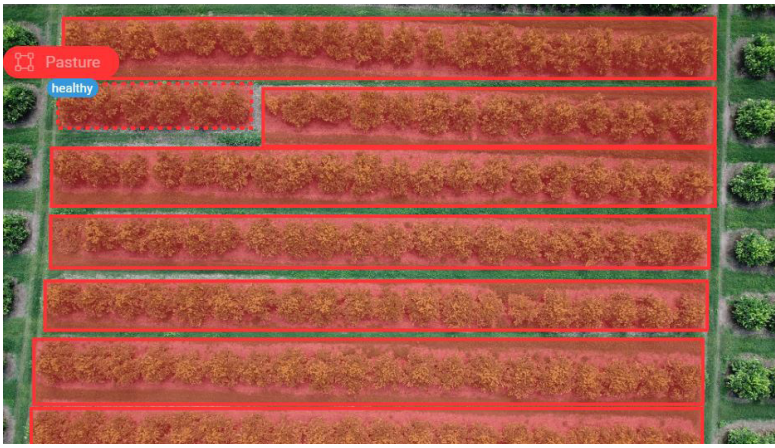
STEP #4
Model automatically detects areas of interest

STEP #5
Data Managers can now zoom into relevant regions

This process takes the same model, and feeds it with the newly discovered symptoms, allowing their annotators to better classify troubling symptoms or infestations. This process is done as part of SeeTree's day-to-day operation and delivery to the end customer.

What about model performance?

The job of the Data Scientist at SeeTree is to determine the model performance. The Data Scientist will then choose to upload the data to the Dataloop platform, or not. In a certain grove, if the model didn't perform well, then this means that the model didn't classify the trees correctly. The Data Scientist will then upload the specific data items using the SDK in order to annotate those trees correctly and add them to the existing dataset, which will later be fed back into the model. SeeTree continuously monitors and updates their findings in reports to which the farmer has access to, and acts on this newly acquired data.

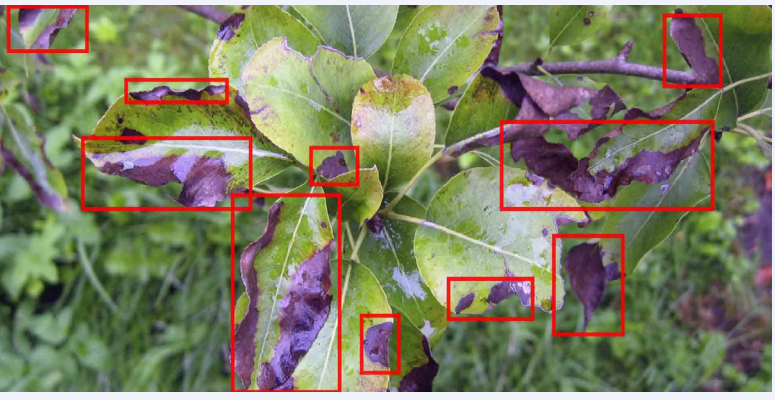


This allows farmers to stay ahead of threats and act much quicker and potentially root out harmful damage.

Without automation, scaling becomes very challenging, time-consuming, and expensive.

Implementing the "Golden Set" to Assure Annotator Accuracy

When it comes to identifying infected or diseased trees the "symptoms" or attributes may be difficult to identify, or scarce. This is where defining the ontology of your dataset will become the building block to your model and will help you define the object detection your trained model provides. You'll also be able to duplicate a "Golden Set" dataset without annotations and assign it to an annotator you wish to test. Then, compare the annotator's results with the golden set to check the annotator's accuracy. SeeTree uses these annotations to train models in order to detect all types of symptoms on the trees.



Or it could be detecting dry trees, which could indicate the tree is dying. But this is not always clear cut, sometimes the attributes are a mixture. When creating a labeling task with Dataloop, you can separate specific labels. For example, "active and green" or "only dry", or "both".

Implementing the golden truth

It's easy to get confused. Therefore, SeeTree utilizes Dataloop's annotation assessment in order to ensure annotator accuracy, this is called the "Golden Set" (ground truth). At SeeTree, the annotator's integrity is very important to them. To be an annotator for SeeTree, you have to do a great job, identifying and classifying diseases, and with the help of the "Golden Set", this is possible.

Conclusion

At SeeTree they've managed to reinvent farming. They've taken the boots on the ground concept,and the farmers' initial problem of not having enough eyes and introduced AI to the picture. This allowed farmers to get as close as they can to actually being on the field, at the fraction of the cost, in the fraction of the time.

Dataloop provides SeeTree's farmers with actionable insights in real-time that brings farming to a whole new level.

"Dataloop and its different interfaces - GUI and SDK - enabled us to embed a sophisticated labeling process, continuously improving our production models on a weekly basis in a variety of classification problems. The team at Dataloop helped us connect our systems in a very efficient way so we can enjoy SeeTree's ability to extract pixels for every single tree and then detect symptoms from those pixels."

Guy Morgenstern, Co-Founder & CTO at SeeTree