

## Canopy Cover Evaluation Through Image Analysis

By Connor Osgood

As the cost of computer disk storage space drops and CPU processing speeds increase, data collection through image analysis has become an ever-important tool in the agricultural researcher's toolkit. Plant *phenotyping* concerns the application of technology to quantify differences in plant traits not easily captured through yield or other metrics. While still a nascent scientific field, plant phenotyping relies heavily upon image analysis for numerous assessments.

The adage follows that "a picture is worth a thousand words", however in the context of scientific image analysis this is an understatement. Many modern smartphones are equipped with 16 megapixel (MP) cameras. With a pixel resolution of 4920 x 3264, a single photo taken by a 16MP camera contains over 16 million individual pixels. If this digital image is thought of as a data table, where each pixel is an individual point representing values such as hue or brightness then it is arguably more accurate to say that "a picture is worth 16 million data points".

Image analysis is especially valuable to agricultural field research because it allows for objective evaluations of plant growth traits recognized intuitively by the human eye, but often interpreted subjectively.

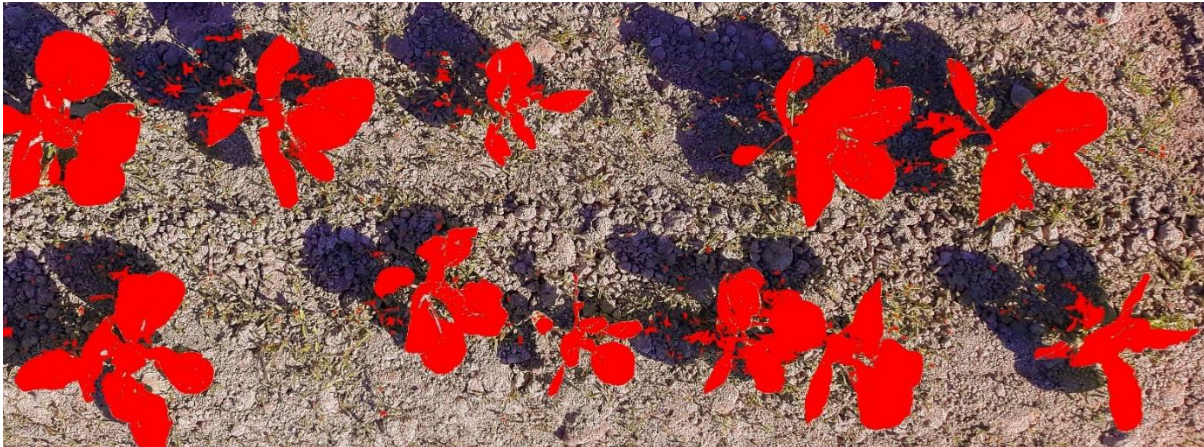
For example, a researcher inspecting a plot of cabbage plants may recognize that one group of plants is obviously larger than another. While a clear difference may be intuitive, how exactly would one measure it? It would be possible for the researcher to take several metrics for each plant including plant height, diameter etc., but this becomes extremely time consuming, and the accuracy of measurement may "drift" as the worker tires.



*Figure 1. Overhead image of young cabbage plants*

Image analysis overcomes these challenges by removing bias from the evaluation. By taking standardized images of each plot at a consistent height and lighting conditions, the canopy cover of the plants can be evaluated *in silico*, in the comfort of an office.

The basic principle that image analysis takes advantage of is that pixels of interest (i.e. the pixels representing the plants) can be isolated based on one or more values. In a well-photographed plot, it can be assumed that green pixels represent plant material, and that the remaining pixels represent the soil surface, debris, irrigation tubing or other components of the photo that are not of interest. From this, the appropriate cutoff values are established to define which pixels are green 'plant pixels' and which are background noise. Once these values are assigned, the software 'lifts' the relevant pixels from the image and subtracts the background, leaving an image containing only plants. This process is known as *thresholding* and is the key component for successful image evaluation.



*Figure 2. Original image thresholded. Red areas represent the green plant pixels detected by the software. Note the presence of noise, smaller clusters of red mainly in shadow areas. These are 'false positives'.*

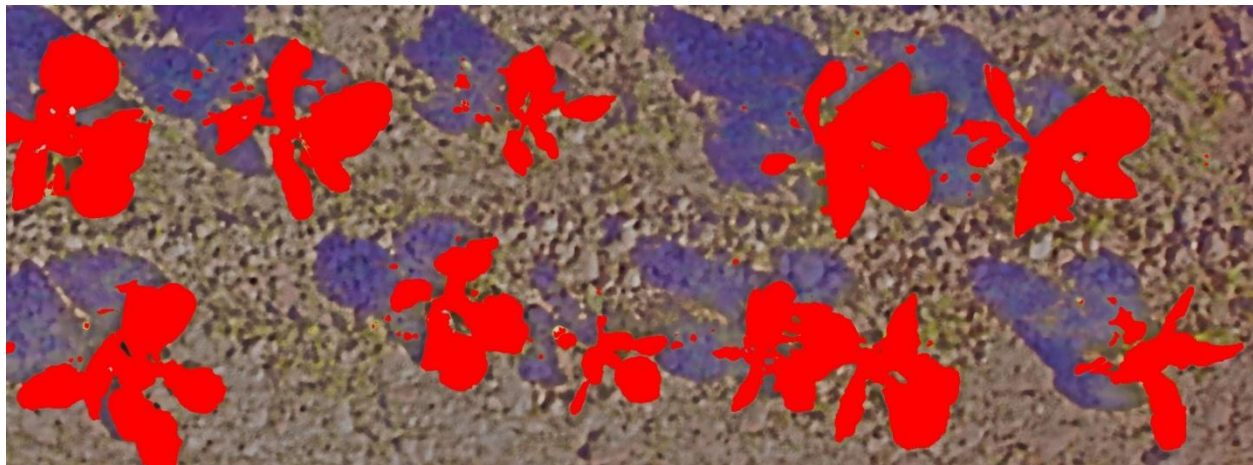
The field images are first cataloged into the workstation computer, so that it is clear which photo belongs to which plot. This is especially critical in small plot research where each plot is often evaluated in its entirety, as opposed to taking several subsamples of an otherwise homogenous field, making the case that identity preservation is critical. Before image thresholding a be performed, the images sometimes need to be pre-processed to compensate for noise, specular light from intense sunlight and other real-world variables. Usually, Guassian blurring or employing a median or Kuwahara filter can eliminate noise to allow for clean isolation of the plants, without losing their overall size and shape. In some cases, other filters may be used including frequency domain filters which can help reduce shadows or glare or enhance ambiguous edges.

While image analysis software packages are equipped with many powerful and purpose-built image processing functions, it is best practice to make every effort to capture an image that does not require heavy processing in the first place. The mantra "garbage in, garbage out" is especially relevant. Mathematics and, increasingly, artificial intelligence, can compensate for some image quality issues but simply cannot make up for detail that is lost in the original image. Image pre-processing should be thought of as an enhancement action, as opposed to a reparative one. While this may seem common sense, there are several challenges that make capturing 'perfect' images in the field challenging. These include hard shadows, reflection on leaf surfaces, wind, unwanted plants (i.e., weeds) and so on.

From the outset, images are taken at the same height, angle and orientation every time. This reduces the likelihood of capturing unwanted elements in a photo such as glimpses of neighboring plots, technicians' boots etc. Additionally, standardizing capture height and angle ensures uniformity among all sample images. If image A is high-quality, then all other images in the sequence should have the same quality. If unwanted noise is present in image B, then at least the resulting error during analysis is uniform plot-to-plot, and an unbiased comparison can still be made.

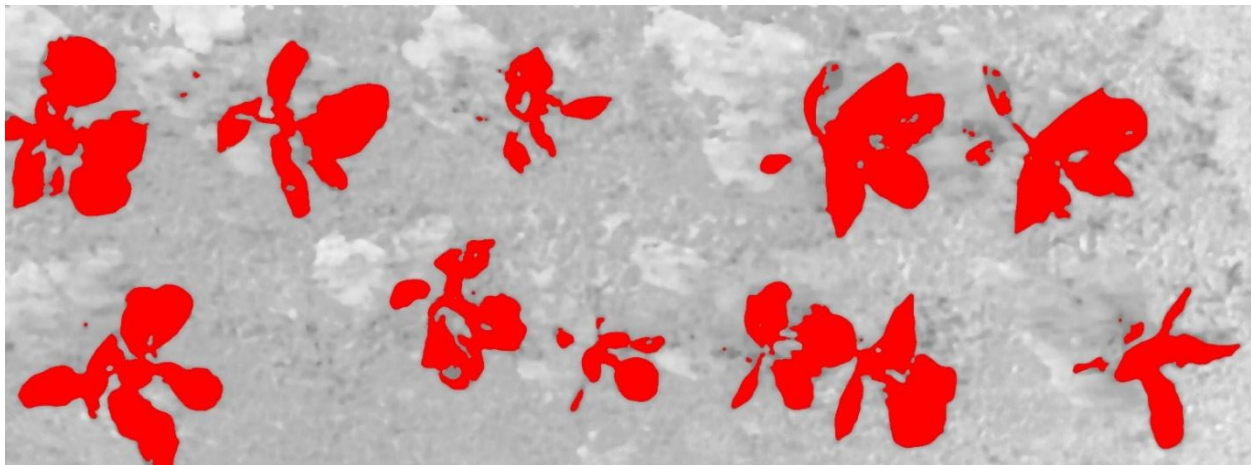


*Figure 3. Original image with median filter applied. Notice 'simplification' of colors, and loss of fine details. This helps reduce noise in thresholded image.*



*Figure 4. Cabbage plot image with both a frequency domain and median filter applied, then thresholded. Noise is reduced in the threshold mask.*

Alternatively, digital images can be processed and interpreted in a multitude of color representations, by translating them to a different *color space*. For most purposes, a digital image is captured from the camera's raw sensor data, and ultimately displayed in red-green-blue (RGB) color space. As the name implies, the image is created by mixing red, green and blue to obtain the final colors. In a sense, an RGB image can be thought of as a composite image of three greyscale images, each representing a color channel (red, green and blue). This property is useful, in that one of the three channels may be the only one that contains relevant color information, and the other two can be deleted. For example, if one were to analyze an image using only the green channel, the image would be represented as a greyscale image with the black and white values representing the absence or presence of green in the composite image. Many other color spaces exist for a wide variety of applications ranging from printing on paper or fabric to television broadcasting. Often, CIELAB color space yields good thresholding results by working off the 'a\*' channel, which represents the red/green balance of the image.



*Figure 5. Image thresholding performed on the a\* channel (red/green) of the source image, after conversion to CIELAB space and applying a median filter.*

It is equally important during image analysis to apply each cleanup and/or analysis step uniformly to each plot photo. Thankfully, software is used to automate these steps so that every photo is treated fairly, also saving the researcher hours of manual work.

After the plot images have been carefully captured from the field, brought into the software and cleaned and thresholded, they are finally ready for analysis. Stand count and canopy cover are straightforward metrics that can be calculated from overhead plot photos.

Stand count is a useful metric for stress trials; if an experimental product is expected to have stress-reduction effects then it is relevant to take stand counts at several points after transplant or seed emergence to calculate percent mortality. Certainly, stand counting can be done simply by hand, image analysis reduces errors through double-counting or missing plants. Additionally, the photos taken are valuable snapshots as they can be revisited later for reference.

Canopy cover is measured on a percent-area basis and is valuable for evaluating growth responses. Generally, each plot in a trial is imaged at various points of growth. For example, imaging would be conducted at transplant, 3 days after application (DAA), 7 DAA, 21 DAA, etc.

The images are thresholded, and the area of green pixels of representing plants is divided by the total area of the photograph, thus yielding percent canopy cover. In other words, the analysis is computing what portion of the image is occupied by plants. After canopy cover is calculated for each treatment at each evaluation date, the *change* over time is compared. If product A increased in percent canopy cover at a faster rate than product B, this would indicate that product A has some influence on the vigor or growth rate of the plants.

Another advantage of canopy cover is that it is a *relative* metric, as opposed to an *absolute* one. It does not require actual area measurements (e.g., cm<sup>2</sup> of leaf area) to complete the assessment, although this is still possible. Being a relative metric, it is easily used to compare the differences in treatments, which is a goal of running a product testing trial.

Given that in-field phenotyping is an emerging technology that is rapidly evolving, the ability to capture many other metrics is also being investigated including mapping plant morphological characteristics such as leaf angle. At present however, many of these applications are still developmental or are highly specialized, whereas overhead plot photography serves as a simple yet reliable technique for assessing growth differences.

See RD4AG's White Paper "Using NDVI for BioStimulant Assessment" by Fabian Sanchez for a look at how specific light spectrums such as NDVI can be used in plant analysis.