filter

After acquiring the image from datasets, the image needs to be filtered and preprocessed first ensuring noise reduced. Since the noise is a high-frequency signal, the Gaussian filter and the median filter are quite useful in this task to remove that. The transfer function of the Gaussian filter is in the form of equation (1), and Median filter shown in equation (2).

Color format of Images

Normalized RGB format

Considering that the color of plants we tried to segment is green, which is also used as a feature in the RGB color space. In this case, the image can be divided into green and non-green parts by using a threshold method. The segmentation of plants can then be achieved. In the RGB space, the color of any pixel can be seen as a combination of red, green, and blue components in different dimensions. An image is a combination of three gray levels of red, green, and blue. If those three components at a certain point are the maximum value of gray, the pixel shows as white color.

However, since the RGB color space is greatly affected by the illumination, in order to reduce the influence of it, the normalization process is first performed in the RGB space. That is, the green component g, red component r, and blue component b are extracted from the image, which are normalized to R, G, and B (see equation (2)). In order to distinguish the green characteristics, consider using the [-1　2　-1] mask template to process the channels (R, G, B) at each point of the image, as shown in equation (3). The template emphasizes the green component, suppresses the red and blue components, and the template sum is zero. Which makes it symmetrical, and the length is odd, so it has a linear phase, which guarantees the time-invariant requirement.

Ideally, for the areas where blue and red are prominent, the result of using the template operation must be negative, while the result of the white area applied the template operation is 0, and only the green area shows a result greater than 0 after the template operation. The calculation result of other color area templates must be negative. This feature is very important. As long as setting value above 0 as the basis for thresholding, the green components could be easily segmented, avoiding the choice of threshold in other algorithms, also reflecting the Irrelevance of algorithm-influence on the image. Since addition and multiplication are used for each pixel from normalization to template operation, the algorithm can be efficiently implemented on the digital signal processor (DSP). While ensuring the performance of the DSP may also achieve high-speed processing to meet the requirements for processing speed in a high throughput environment.

2 Image segmentation based on HSV color format

The RGB color space applies a linear combination of three color channels to represent colors. Any represented color is related to these three components, and they are highly correlated. Therefore, it is not intuitive to continuously change colors. Adjustments need to be implemted on that.

By observing the input data of this task, it can be found that the natural lighting, occlusion, and shadows in the photo make a greater impact on the brightness of different parts of the picture.

In the RGB color space, all three color components are closely related to brightness. In other words, as long as the brightness changes, they will change accordingly, and there is no more intuitive way to express a same channel. Since all image processing approaches should be accessed by human eyes. In monochrome, eyes are the least sensitive to red and the most sensitive to blue, so the RGB color space is a color space with poor uniformity. If the similarity of colors is directly measured by Euclidean distance, the result will have a large deviation from human vision. For a certain color, it is difficult to infer a more accurate three-component value to represent.

Based on the above reasons, the HSV color space is more commonly used in image processing, which is closer to human perception than RGB. It is very intuitive to express the hue, vividness and brightness of the color, which is convenient for color contrast. In HSV color space, it is easier to track objects of a certain color than BGR, and it is often used to segment objects of specified color.

For these tasks, there will be a big difference between the leaves of the plants and the background traies. The direct difference of the plant leaves is only reflected in the color tone, brightness and saturation. For further identification of leaf area grown condition and maturity also depends on these key parameters. In this project, there are some horizontally compared results between two different color space formats regarding the accuracy of identification as well as prediction.

Morphological processing

The image processed by the above methods will still have some stray points shown. At this time, the open operation in the morphological algorithm can be used to remove [10]:

The opening operation of the structural element B on the image A is denoted as A $\circ$ B, as shown in formula (4), the image is first corroded and then expanded. Corrosion could help to remove stray points, but also make the edge area smaller, so it needs to be restored by expansion.

Corrosion is denoted as AB, as shown in formula (5), where A is the complement of A, and is the empty set Θ .

The expansion is denoted as AB, as shown in equation (6), where B is the symmetrical image of B about the origin, which is an empty set.

Repeat the process until it satisﬁes the tolerance or error value.

DSC:

DSC describes the similarity between two contours by estimating the percentage of the volume (area) of the overlapped part of the two contours to their total volume (area). The calculation formula is:

Among the formula, A and B respectively represent two target areas; \union represents the common part of the two; + represents the sum of the two. This calculation method is the most used comparison method [8]. In fact, DSC is a special case of Kappa coefficient [9]. In extreme cases, the calculation formula of Kappa coefficient can be simplified into the current DSC calculation formula. As early as 1945, Dice [10] mentioned this parameter when studying the number of different species in ecology.

The DSC calculated by equation (1) is a parameter with a value ranging from 0 to 1. When the value reaches 1, it means that the two target areas are completely overlapped, and when the value gets 0, it means that the two target areas are completely separated.

DSC is very sensitive to differences in position and size. For example, if two graphics of equal size overlap half of their areas, they only have half of each other, resulting in DSC=0.5; or one area is completely covered by the other half smaller than it In the upper area, DSC is only two-thirds. In the process of DSC measuring image similarity, the difference in position can reflect the difference more than the difference in size. For images with the same volume (area), different positions will cause the final DSC to be different. This can reflect the intuitive feeling that two areas (one of which completely contains the other area) are more similar than two partially overlapping areas. Combined with the actual contour image processing, although DSC has the advantages of conciseness and clarity in the calculation process, it cannot describe all situations.

Text

Description automatically generated

­­­A picture containing pot, flower, drawing

Description automatically generated

A picture containing table, lit, sitting, computer

Description automatically generated

A picture containing table, lit, sitting, computer

Description automatically generated

A picture containing pot, flower, drawing

Description automatically generated

RGB

Text, application, chat or text message

Description automatically generated

the main purpose of this task is to calculate the Plant Leaf Area, which should be achieved through image segmentation.

First is preprocessing,

which covers color formation ,

noise reduction.

For color formation,

here compares two different formats:

Normalized RGB and HSV

We found that

compared with HSV,

the final processed image of Normalized RGB got some details lost, and the metrics shown in table also approved that.

For noise reduction part,

we have applied the opening operation\ closing operation

from morphological algorithm with the combination of Gaussian Filter plus Median Filter.

A sample processed image shows here.

After pre-processing, we have implemented segmentation with

watershed

\ Kmeans clustering

and thresholding without other algorithms.

For the K-means clustering function, the input K actually is the number of drawing boxes from task1.

As results, The Gray table are the metrics of these three segmentation methods.

As It shows, the watershed and K-means performed better.

for the Dice Similarity Coefficient, K-means gets higher scores, and watershed gets higher in Intersection Over Union,

For the reasons

When we have an inspection on the

best performance image and

the worst,

I believe it is because Dice Similarity Coefficient concentrate more in the True Positive part instead of Intersection Over Union

And K-means performs better on removing the noisy item,

Which could be seen at the right bottom,

K-means removed the edge of tray but other methods could not make it.

In addition, iterations K-means iterations blurred edges of leaves.

In conclusion, when the target datasets is not too large,

K-means performs better, otherwise watershed segmentation should be considered.

Some tasks like task3, which pursuit further segmentation with a large dataset, we recommend MASK-R-CNN,

Please let my mate Gewei to introduce that.

Kmeans 需要循环而将植株叶片边缘模糊了，