Detection

Research

Name: Sieuwe Elferink
Class: T31 Delta
Date: 18-11-2019

Introduction

After having finished the previous research about what the parameters are to know if a beehive is healthy or not, I found two parameters which are interesting. These are the Brood pattern state and the amount of bees inside the beehive. These are interesting because they are both measurable with modern state of the art technologies and they have a large impact on the health of the beehive.

Now that I now which parameters to detect, there has to come a method to detect the parameters. The different methods for each parameter will be written down below. At the end there will be a conclusion comparing all methods and decide which is best for each parameter.

Results

Clarification

All detection methods explained below are image based and not video based. This is because the health of the beehive does not change per minute and not even per day. That is why instead of analyzing a video which takes multiple images per second, I will make one high resolution picture per day. This picture can than be analyzed and then store the result to give a trend. This is not only decreases the computational power needed it also makes it easier to get high resolution pictures with lots of detail because the camera can take the time to let the light shine into the sensor the get a high detail image.

Brood Pattern classification

1 – Using traditional CV methods

   - Explanation
   To detect the brood pattern we can use CV (Computer Vision) by using a Camera. Then using OpenCV I can analyze the incoming Image. To detect the brood pattern in the frame I can apply a HSV filter which filters out all other cells and only shows the brood cells based on color. After I can detect the contours of the available cells. Based on the contours I can detect the brood pattern cells locations. After that I can loop over every brood cell and detect the brightness of every cell. If the brightness is low the cap is probably off the brood cell. If the brightness is high the cap is probably still on the brood cell. Based on how much caps are off the total amount of cells I can classify the entire brood pattern in healthy or not healthy.

   -Advantages
   It is very fast and does not need an special dedicated GPU for computation.

   It can be very accurate if the detection works well. And it can even say how much brood cells are healthy and how many brood cells are broken.

   - Disadvantages
   If this way of classification is even possible it will be vary inaccurate because it is based on color. This is because color can change overtime and is dependent on light. Also I am saying "if this way is even possible". This is because applying a Hue filter for colors which are so close together are very difficult and ins ome cases even impossible. And even if there is a correct Hue filter it will be very difficult to detect the contours for the cells because they are very small, Only a couple pixels per cell.

2- Using a CNN

- Explanation
Another way of detecting and classifying the brood pattern is by using an CNN (Convectional Neural Network) which is trained on a large data set to learn shape and edge detection like Inception v3. We can then download this trained model and then use transfer learning to retrain the last layer where the actual classification happens. This retraining is done with our own data set of healthy brood patterns and unhealthy brood patterns. After the training is done the CNN has learned which features of the image belong to which class and can then classify weather the beehive is healthy or unhealthy and give a percentage of how sure it is.

-Advantages
It is more reliable in different situation because it is not based on only color but on hundreds of abstract features which we humans can not see but the CNN can. Especially if the training set contains images from different angels, times and weather conditions.

It is easier to implement because the pretrained model can be easily downloaded. After gathering a dataset the model only has to be trained with some easy code and after that the model can accept images and give the class.

- Disadvantages
It is slower than a CV based detection because a CNN requires a lot of small calculations because the CNN consists of hundreds of thousands of neurons with weights between them which have to be calculated every classification run.

Because of the above mentioned disadvantage the CNN needs a expensive GPU for computation. This is because the CNN has to perform a lot of small calculations besides each other.

A CNN needs a large high quality dataset to be trained on. Getting hundreds of images of the inside of beehives which show the correct parts of the beehive may be very difficult because not a lot of beekeepers make pictures of their beehive.

Population counting
1 – Using traditional CV methods

- Explanation

The amount of bees in an beehive can be counted using traditional CV methods. This because I don't have to count every single bee. If I can make a approximation for the amount of bees and then can show a down worth trend, I can tell that the beehive is getting unhealthy. This is because if there are is a decrease of 10 bees in a day it says nothing about the beehive health. Only if there is a trend of a lot of bees disappearing it says something about the health. Because of this it also does not matter if the count is 50 bees off the actual count. Because if the count is always 50 bees off it still shows the down worth trend. And even if the count is one day 20 bees to much and another day 20 bees to low it does not matter. Since there has to be a trend over multiple days that bees are disappearing. So if the next day shows a increase again the average will be that not a large amount of bees has disappeared.

To detect the population I will use a camera pointed to the beehive. Because all the bees have a very distinct color compared to the hive grid, I can apply a HSV filter. This filter will filter out all the background and only shows the cluster of bees. After the filter is applied I can calculate the contours of the bee cluster. With the contours I can calculate the area the bees cover in pixels. The more pixels the bee cluster covers the more bees there are and thus the healthier the beehive is.

-Advantages

This detection method is robust in the sense that it does not matter that there are bees covered by other bees what happens a lot because a beehive is very crowded and bees walk over each other.

This method is fast and does not require an expensive GPU.

- Disadvantages

This method does not give an exact bee count. In only gives an approximation.

It may not always be reliable because it is based on color.

2- Using a CNN

- Explanation

Another possibility is to use a CNN to count the bees. In this method we will transfer an existing model to recognize a bee. A camera placed inside the beehive gives a image. This image will then be split into smaller images to make it easier for the CNN. Every splitted section of the image is then analyzed by the CNN which will give a list of every single bee with its location. We can then get the length of this list to know how many bees are in the beehive.

-Advantages

A CNN will give a much more accurate list of bees because it counts every single bee inside the beehive.

A CNN will give the exact location and size of every bee. This can be useful for other applications like classifying the health for every bee.

- Disadvantages

A CNN will require a large data set which sometimes is difficult to find. (In this case there is a free open source data set available).

A CNN needs a lot of computational power and thus requires an expensive GPU.

The detection will not be accurate since bees are very small. Because CNN image classifiers are known to find small subjects difficult to detect it will miss a lot of bees. The CNN will also miss a lot of bees because it will not detect a bee as a bee if it is covered by another bee.

3 – Using electrical counter

A electrical counter placed at the entrance like an IR array which counts how many bees fly in or out. While this is computation inexpensive it does not say a lot about the health of the beehive. This is because weather is an important factor for bees to fly out or stay inside. This can result in the computer thinking that the health of the beehive is lower because it is actual the same but only the weather not favorable. Also the amount of bees flying in and out of the beehive does not say a lot about the health of the beehive.

Bees health classification
1- using pre made solution

- Explanation
During this research I found an already implemented bee health classifier based on a CNN. This CNN is trained on hundreds of pictured of bees while flying in different situations. A camera is placed in front of the entrance. Every time bee flies past this camera the CNN will recognize it and classify if the bee is healthy or not.

-Advantages
It is free and open source

The entire program is made in python.

All the development work is done so it only has to be started to work. And there are instructions on how to do this.

-Disadvantages
The program relies on a camera placed in front of the entrance of the hive while the other two parameters relies on camera placed inside the hive. Because of this two camera's have to be placed which result in higher cost and complexity.

Conclusion

Brood Pattern classification
To recognize the brood pattern the best method is a CNN. This is because a brood pattern does not have a single main visual feature which is very recognizable compared to the other parts of the beehive. This makes it difficult to make a program by hand which extracts the needed features to classify the image. Instead it is better to use a CNN which automatically extracts this features and classifies the image. The only downside is that a large data set is needed. Because there are no available data set, I have to make my own data set which will take some time.

The computational disadvantage of a CNN does not matter since there will be already a GPU in the PC because other parts of the beehive project will also need a GPU.

Population counting
To count the amount of bees a traditional CV method is the best because it will be the most robust. A CNN will have a lot of false positives or miss a lot of detection's because it has to detect every single bee. Also if a bee is on top of another bee the CNN wont detect the bee below as a bee. Also because it is not necessary to detect every single bee because of the reasons explained earlier it will add a lot of unreliability for a feature which is not needed. Because the CV method will only count the cluster of bees it does not have these problems.

Bees health classification
For the bees health classification there is only realistic option which is to use the pre made one. This is because it is really easy to implement in our beehive project. If I want to make it myself it will cost a lot of time I don't have while I am not even sure that I can make it better than the existing implementation. Because it is so easy to implement it will receive less attention. I will implement it into the project once I Finished the two above mentioned detection methods.