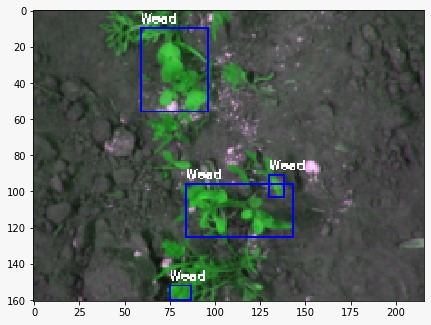


Ministry of Communications and Information Technology

Intensive Training Program (ITI)

Data Science Track

**Weed Detection**

****

**March, 2020**

**Weed Detection**

This documentation submitted as required for the final project

of the ITI nano-degree program

**By**

Omar Hesham Amin

Ahmed Abd Elaty Younis

Menna-tullah Abdelsamie Sharkawy

Powered by

**Teradata**

**Table of Contents**

|  |  |
| --- | --- |
| 1. Abstract 2. Introduction 3. Methods    1. Image Morphological Operations    2. Feature Extractions    3. SVM Model    4. U-Net Architecture 4. Proposed Work    1. Data Preprocessing    2. Image Segmentation    3. Creating Model    4. Prediction 5. Experimental Setup 6. Results and Discussion 7. Conclusion 8. References |  |

1. **Abstract**

Precision agriculture is gaining increasing attention because of the possible reduction of agricultural inputs (e.g., fertilizers and pesticides) that can be obtained by using high-tech equipment, including robots. Information on weed distribution within the field is necessary to implement herbicide application. This research aims to provide a weed detection tool for future agricultural robots. The weed detection tool incorporates the use of machine-learning procedure explicitly implementing U-Net and morphological operations on RGB images to classify crop and weed.

1. **Introduction**

One of the major concerns of an economy is accelerating agricultural development. The term crop defines the cultivated plant, while the term weeds defines unwanted plants that grow spontaneously in the field. Controlling weeds is an important challenging task in agriculture because weeds compete with crops in the field, contributing to a lower crop yield. Weeds are controlled by applying herbicides to the field. We aim to reduce the amount of herbicides without compromising the quality of crops. It can be achieved by selective spraying or accurate mechanical removal of weeds, while achieving that manually is time consuming and expensive.

Automatic weed detection systems can be used to improve the efficiency of precision farming techniques on weed control by modulating herbicide spraying appropriately to the level of weeds infestation. However, the great variety of crop and weeds shapes, size and colors, together with the presence of overlapping between plants, makes automatic crop/weeds classification throughout images a challenging task for autonomous farming robots.

Nevertheless, the capability to generalize the trained models still remains an obstacle to employ farming robots in different farm conditions, e.g., caused by environmental changes, plants characteristics and types of soil.

We have proposed a pipeline to detect and mask the presence of in an image. The image data is pre-processed in order to apply some morphological operation to detect weeds and fed them into UNET model which performs semantic segmentation and generates masks over weeds.

The main contributions of this work are:

* A background removal method that uses a deep pixel-wise segmentation to distinguish between soil and plants.
* Morphological erosion and dilation methods are used to disconnect objects within the image.
* An accurate crop/weed classifier based on U-Net.

The limited amount of data with pixel-wise annotations is in fact the bottleneck for most crop/weed classification methods. Our method relies on generating more data by flipping the images.

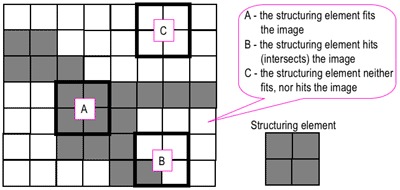
1. **Methods**

For classification of weed images, we have used the morphological operations and feature extraction techniques and we have also used the following classifiers – Dense Layer with Softmax Activation (discussed in the U-Net section) and Support Vector Machines.

* 1. **Morphological Operations**

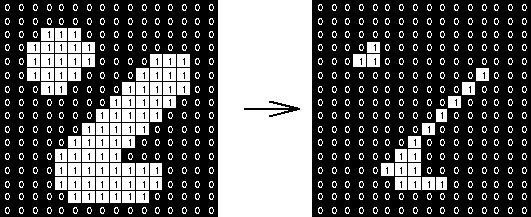
Binary images may contain numerous imperfections. In particular, the binary regions produced by simple thresholding are distorted by noise and texture. Morphological image processing pursues the goals of removing these imperfections by accounting for the form and structure of the image. These techniques can be extended to greyscale images.

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. According to Wikipedia, morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to greyscale images such that their light transfer functions are unknown and therefore their absolute pixel values are of no or minor interest.

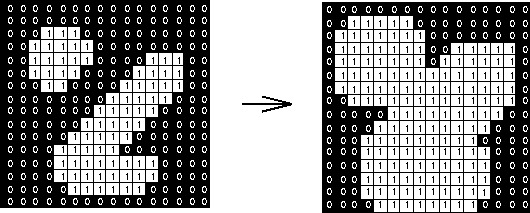
Morphological techniques probe an image with a small shape or template called a structuring element. The structuring element is positioned at all possible locations in the image and it is compared with the corresponding neighborhood of pixels. Some operations test whether the element "fits" within the neighborhood, while others test whether it "hits" or intersects the neighborhood:

A morphological operation on a binary image creates a new binary image in which the pixel has a non-zero value only if the test is successful at that location in the input image.

The erosion of a binary image f by a structuring element s (denoted fөs) produces a new binary image g = fөs with ones in all locations (x,y) of a structuring element's origin at which that structuring element s fits the input image f, i.e. g(x,y) = 1 is s fits f and 0 otherwise, repeating for all pixel coordinates (x,y).



The dilation of an image f by a structuring element s (denoted fөs) produces a new binary image g = fөs with ones in all locations (x,y) of a structuring element's origin at which that structuring element s hits the input image f, i.e. g(x,y) = 1 if s hits f and 0 otherwise, repeating for all pixel coordinates (x,y). Dilation has the opposite effect to erosion -- it adds a layer of pixels to both the inner and outer boundaries of regions.



Morphological Gradient

This is the difference between the dilation and erosion of an image. Because dilation and erosion mostly affect the pixels that are close to the boundary between the foreground and background, their difference generally yields the boundary and thus this is used for edge detection and segmentation tasks. Now, let’s discuss how to implement this using OpenCV-Python.



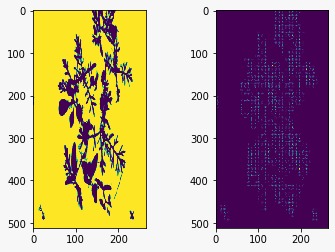
* 1. **Feature Extraction**

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is a large number of variables that require a lot of computing resources to process. Feature extraction is the name for methods that select and /or combine variables into features, effectively reducing the amount of data that must be processed, while still accurately and completely describing the original data set.

* Histogram of Oriented Gradients (HOG) feature descriptor:

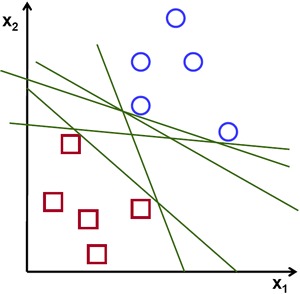
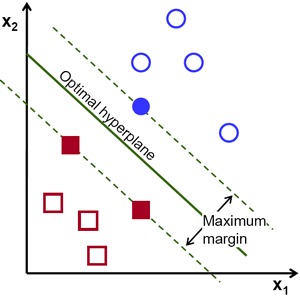
A feature descriptor is a representation of an image or an image patch that simplifies the image by extracting useful information and throwing away extraneous information.

A feature descriptor converts an image of size width x height x 3 (channels) to a feature vector / array of length n. In the case of the HOG feature descriptor, the input image is of size 64 x 128 x 3 and the output feature vector is of length 3780.



* 1. **SVM Model**

The objective of the support vector machine algorithm is to find a hyper-plane in an N-dimensional space (N — the number of features) that distinctly classifies the data points.

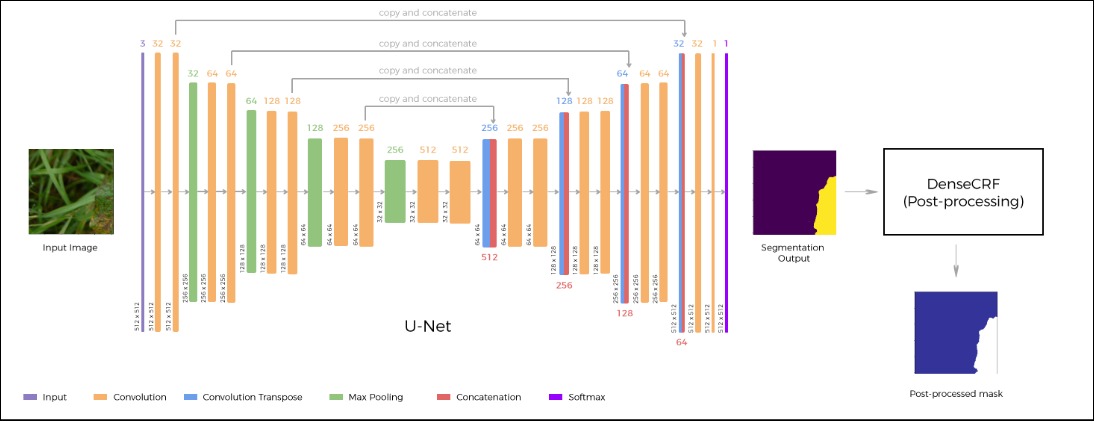


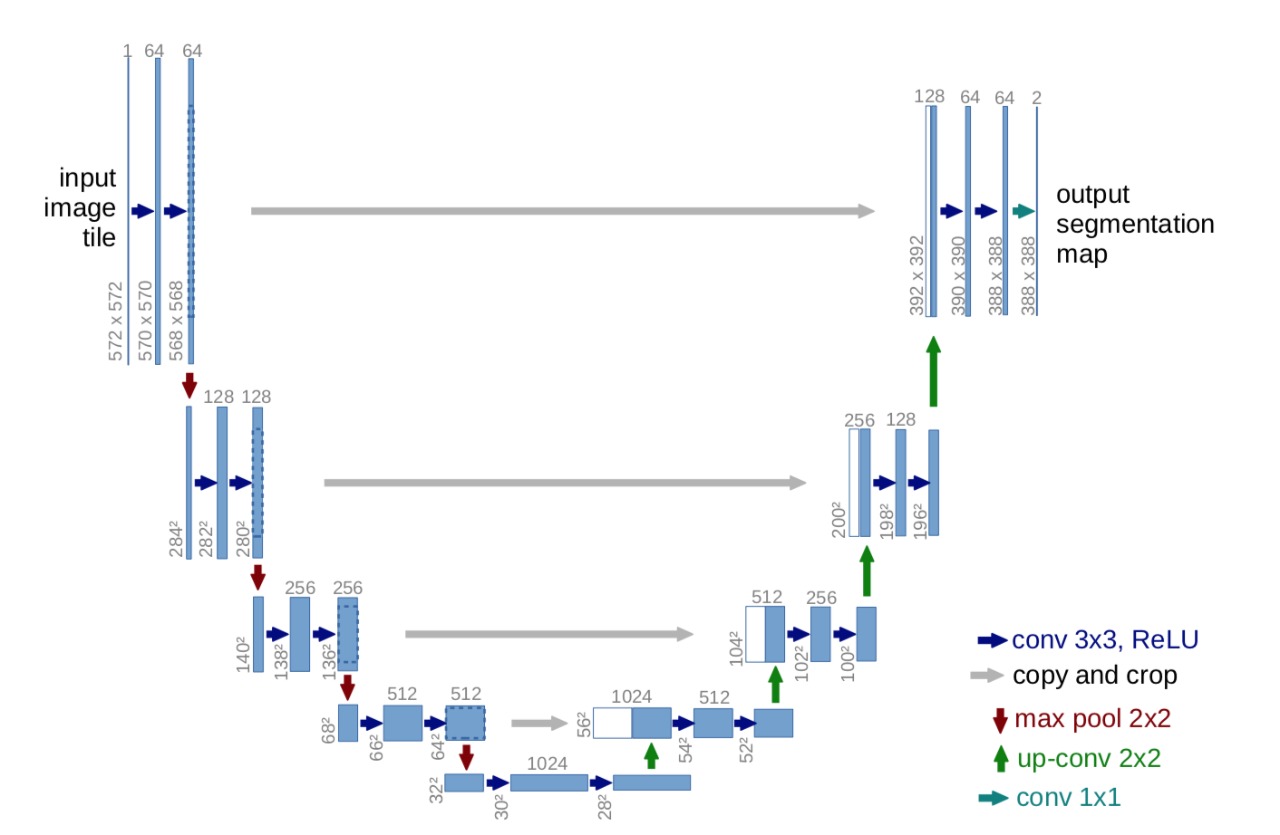
Support vectors are data points that are closer to the hyper-plane and influence the position and orientation of the hyper-plane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyper-plane. These are the points that help us build our SVM.

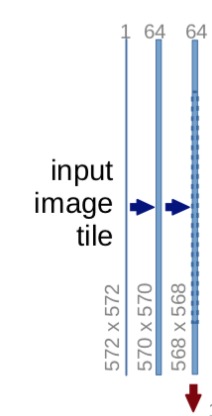
* 1. **U-Net Architecture**

U-Net, evolved from the traditional convolutional neural network.

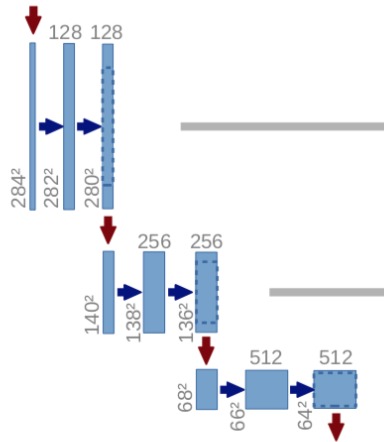
U-Net is dedicated to solving this problem. The reason it is able to localize and distinguish borders is by doing classification on every pixel, so the input and output share the same size.



First sight, it has a “U” shape. The architecture is symmetric and consists of two major parts — the left part is called contracting path, which is constituted by the general convolutional process; the right part is expansive path, which is constituted by transposed 2d convolutional layers(you can think it as an upsampling technic for now).

Notice that each process constitutes two convolutional layers, and the number of channel changes from 1 → 64, as convolution process will increase the depth of the image. The red arrow pointing down is the max pooling process which halves down size of image(the size reduced from 572x572 → 568x568 is due to padding issues, but the implementation here uses padding= “same”)

The process is repeated 3 more times:

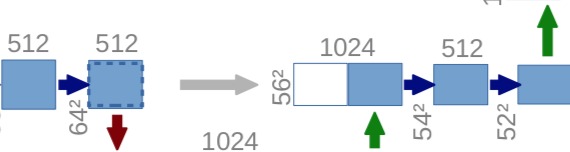


and now we reache at the bottommost:



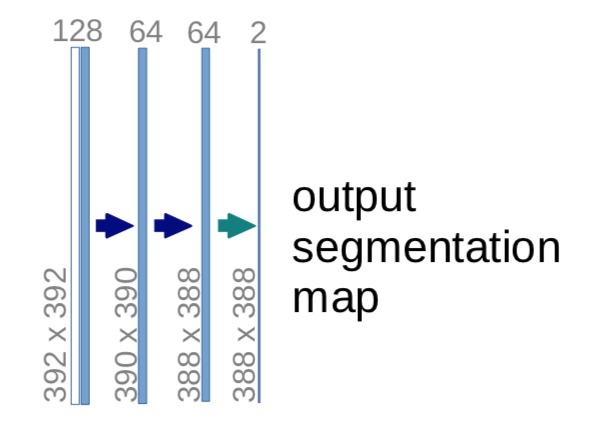
still 2 convolutional layers are built, but with no max pooling. The image at this moment has been resized to 28x28x1024. Now let’s get to the expansive path:

In the expansive path, the image is going to be upsized to its original size.

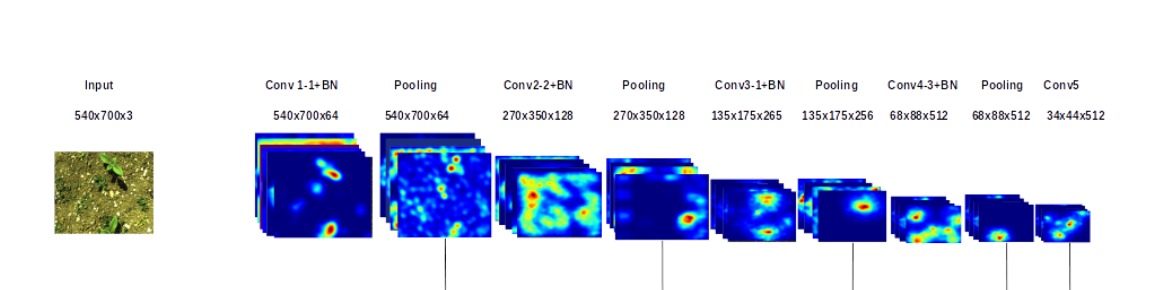


Transposed convolution is an upsampling technic that expands the size of images.

After the transposed convolution, the image is upsized from 28x28x1024 → 56x56x512, and then, this image is concatenated with the corresponding image from the contracting path and together makes an image of size 56x56x1024. The reason here is to combine the information from the previous layers in order to get a more precise prediction.



The last layer is a convolution layer with 1 filter of size 1x1(notice that there is no dense layer in the whole network). And the rest left is the same for neural network training.



1. **Proposed Work**
   1. **Data Preprocessing**

**Loading data set**

Data Set which contains 60 images. The images include weed, crop and the soil. the data is available on GitHub and as we using Google Colab which facilitates loading and reading data and enable storing and loading data from google drive and powered the developers with a very powerful computational and physical power which will discuss deeply in the used tool section.

**Data Augmentation**

The small amount of data and how difficult is the process of collecting data, the time consumed in the process in order to the high cost of taking pictures of different location and circumstances in the field this cause a law accurate of training the models and turning the weights. A popular method to solve this problem without addition efforts or cost is Data Augmentation. Data Augmentation is a technique used to generate more data from existing data by using transformation such as scaling, translation, rotation and Flipping

**Resizing Images**

We resizing images to get a smaller image, which reduces our calculations and in longer run also reduces time complexity and computations

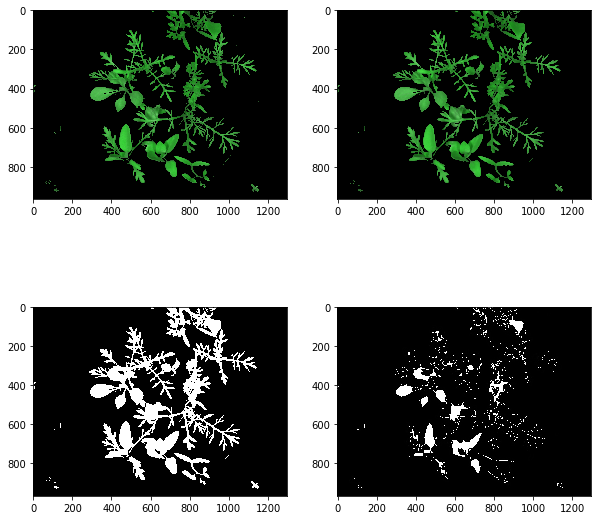
**Create Masks**

Loading the image and convert RGB (red, green, blue) colored image which is based on the color theory that all visible colors can be made using the additive primary colors of red, green, and blue into HSV (which stands for Hue Saturation Value)

Using this model, an object with a certain color can be detected and to reduce the influence of light intensity from the outside.

Creating ROI (Region of Interest) is a portion of an image that you want to filter or operate in some way, Binarization method is based on thresholding to differentiate soil and plant so far we disconnect the foreground from background thresholding is the simplest approach to segment an image which based on setting a threshold value on pixel intensity

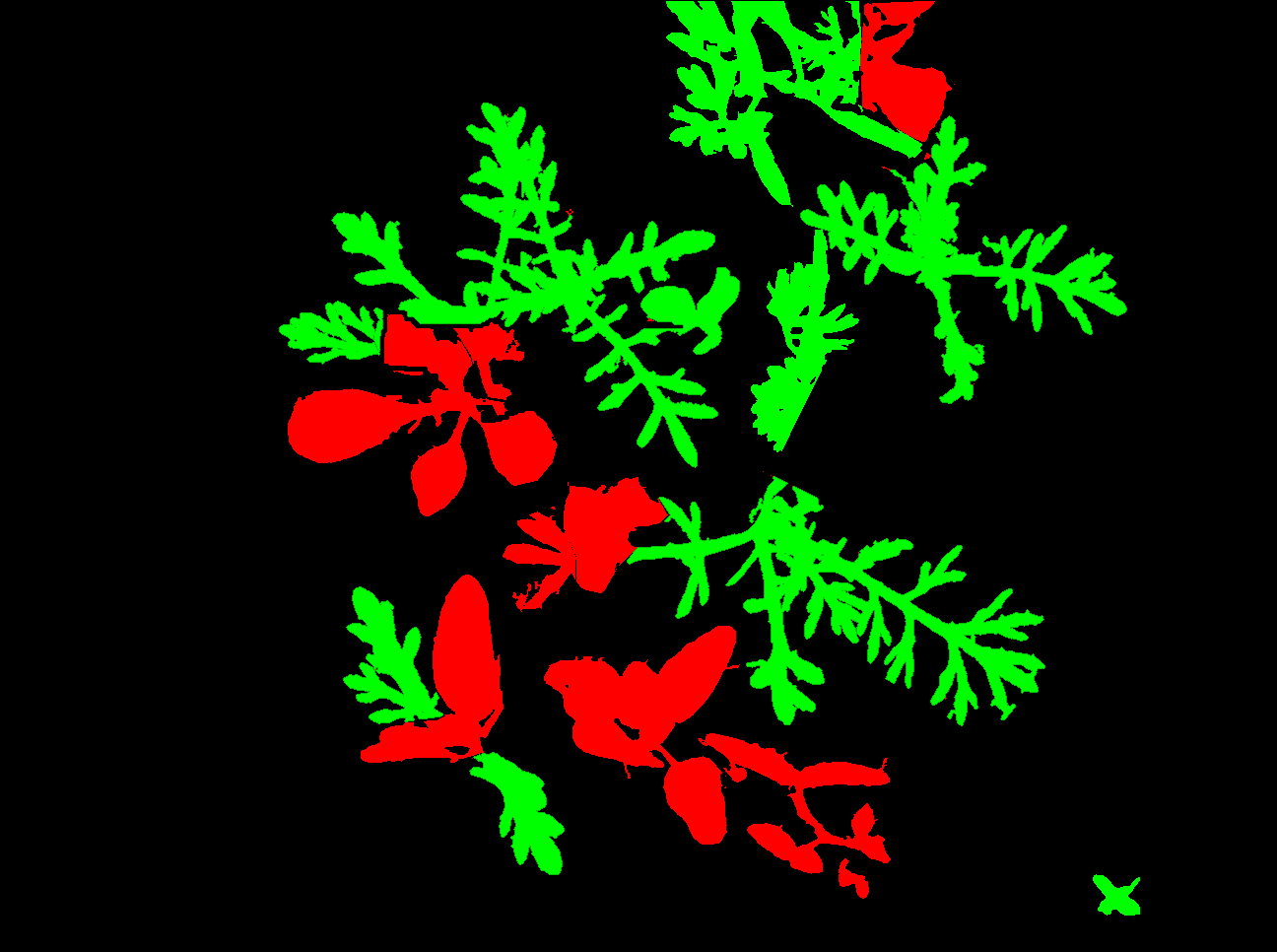
Within our work we use the **Hysteresis** **thresholding** (two thresholds, one at each side of the valley) Pixels above the high threshold are classified as object and below the low threshold as background and Pixels between the low and high thresholds are classified as object

****

* 1. **Image Segmentation**

After sperate the foreground from the background depending on thresholding-based segmentation or colored-based segmentation mentioned earlier in Data preprocessing section the main challenge is how to separate weed from crop or disconnect the plant within the same image which has many cases and still a challenge for the future work. For accurate mapping of weed we use two techniques:

**Hand** **Labelling** is accurate and more efficient way through this process we depend on agriculture experts to separate and distinguish the weed and the crop by coloring the weed with red color and the crop with green color to create our annotations which we used later in in training our model.

****

**Morphological Operation**

After reading the images and pre-processing by applying more than image technique to enhance our images we try to automate the process of labeling the objects within the images by using some morphological operations. We depend on the size of the leaf of the weed which mostly being bigger and lager than the crop leaf by applying erosion and dilation on the binary image and get the difference between the two images we get a disconnected component of each images

****

****

The next step is to find the contours of the remaining disconnected object which is parts of the crops and the coloring these parts to get the weed labeled only the boarders of the weed will be missed because of subtracting the image to get the disconnected component but we final get the shape and the area of the leaves of the weed correctly



This is the general case and the optimal case which is not always available

**Create Labels**

After annotate the images object we use these annotations to create the label to use in process of training the model by comparing the annotated image and the mask get three images and use the in the training phase first the total image



Second one is the weed image which contains weed only



Third one is the crop image it’s contains only the crop of the original image

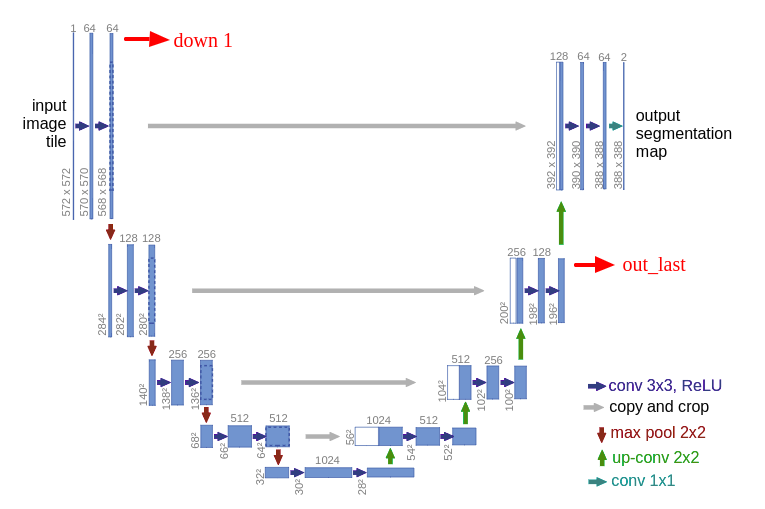


By using the weed image and crop image we get we can train our two models on to detect weeds and another on to detect the crop and also calculate the percentage of the weed in which we could get the highest priority to the part o the field the contains the largest amount of weeds

* 1. **Create Model**

We create two sperate neural network the first on to predict the weed and the second one to predict the crop each on is a UNET architecture

UNET very popular end-to-end encoder-decoder network for semantic segmentation which is Fully Convolution Neural Network used to solve semantic segmentation problem



first part (denoted as DOWN) of the architecture will be familiar. This first part is called down or you may think it as the encoder part where you apply convolution blocks followed by a maxpool downsampling to encode the input image into feature representations at multiple different levels.

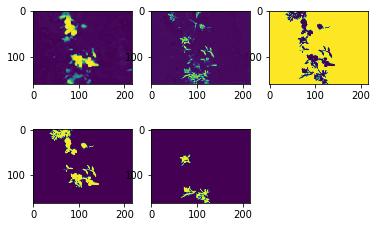
The second part of the network consists of upsample and concatenation followed by regular convolution operations

Upsample are expanding the feature dimensions to meet the same size with the corresponding concatenation blocks from the left The main contribution of U-Net in this sense compared to other fully convolutional segmentation networks is that while upsampling and going deeper in the network we are concatenating the higher resolution features from down part with the upsampled features in order to better localize and learn representations with following convolution. As we discuss earlier the Model is compiled with Adam optimizer and we use binary cross entropy loss function since there are only two classes (salt and no salt)

* 1. **Prediction**

Having two separate models one for each type of planet weed and crop, First is weed model to predict weed and fitted by the original image, the weed image which we created by using the annotations images in the earlier stage

The output of this model is a predicted image of the weed within the image and highlight it in the original image and calculate the percentage of the weed within the input image Second one is crop model to predict the crop we focus in predicting weed which is the real challenge trough our work our model achieve accuracy up to 77% which is acceptable as the amount of data is very small is only 60 image and in addition to the very complicated concern which is the overlapping between the weed and crop in the image and the size of the weed and the crop.



1. **Experimental Setup**

Step1 loading the image.

Step2 Convert image in HSV color space then get binary image.

Step3 Get the mask and then and apply the morphological operation to get the connected object.

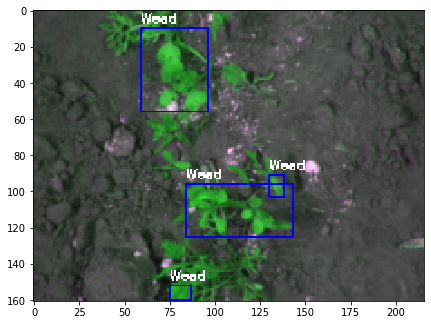
Step4 Coloring the weed and crop.

Step5 Use these annotations to get separated image with the weed and the crop from the same image.

Step6 After creating the model the training step depend on the annotations of each weed and crop as the label for the weed model and the crop model.

Step7 Prediction step is to use each model weed model and crop model to predict the weed and crop in an image.

The final output Is the original image with localize the weed within the image and the percentage of the weed in the predicted in the image.



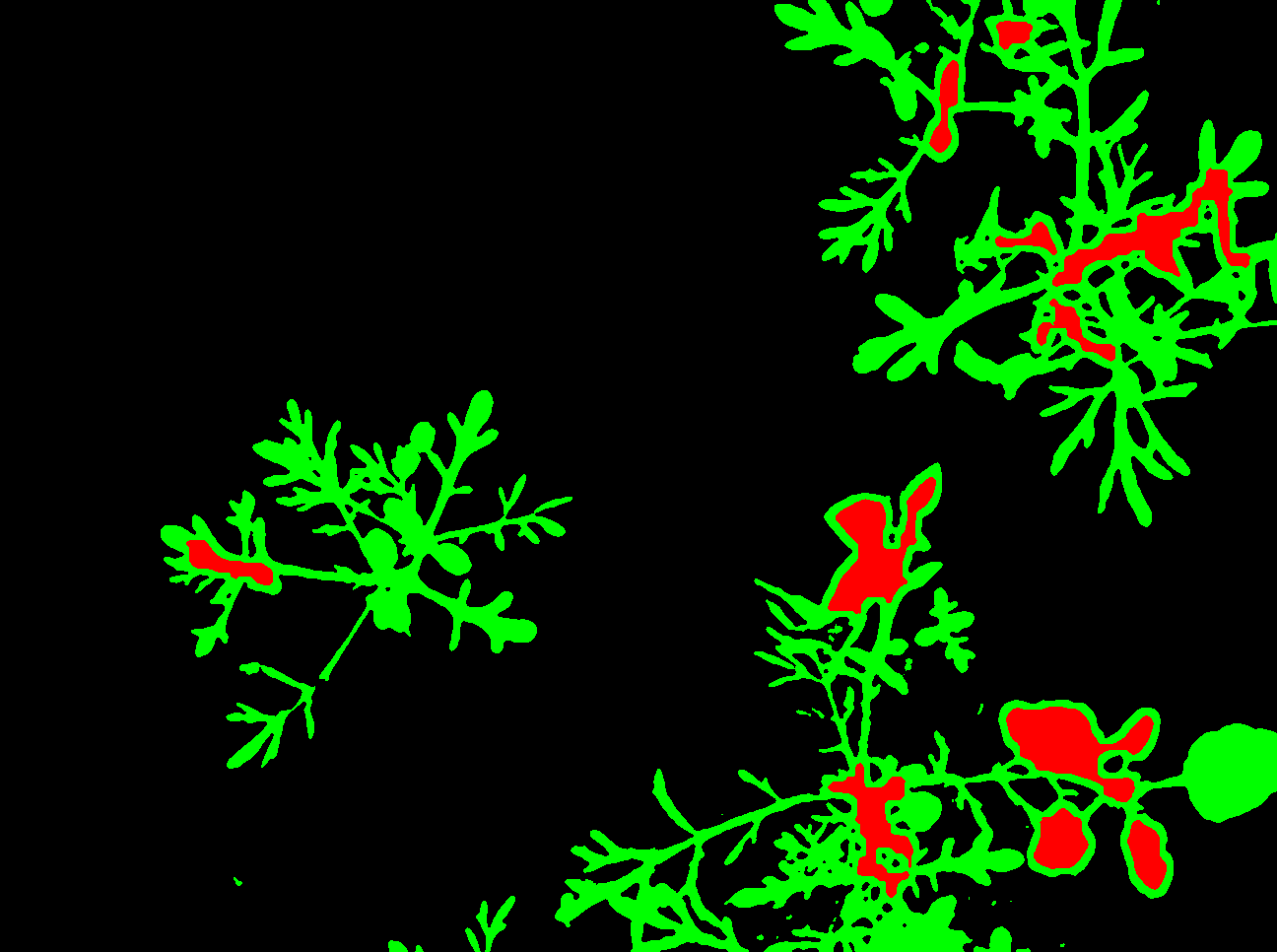
1. **Result and Discussion**

Our work present to main algorithm or strategies as a solution to the weed detection problem First segment and disconnect the weed and the crop in the image and second to use UNET as semantic segmentation to determine the weed within the image, how accurate are these two steps accomplished is based on the position, size and the circumstances on which the image capture from the field and a main condition that has a great effect on the two stage is the overlapping between plant in the images

**Non overlapping** the leaves of the weed and the crop are close but there is no intersection between the leaves of plant weed or crops it make it easy to classify and use morphological operation for determine weed from crop.

****

**Overlapping** include the situation the leave of the crop and the weed are too close and intersected or even the leave of the crop intersected that directly affect in the way the crop and weed separated as it make the shape of the leave for each of them hard to prominent and difficult to get and compare the leaf features and it affect the accuracy and detection.

****

1. **Conclusion**

Accurate mapping of weeds is a pre-requisite for weed density estimations and variable rate herbicide prescription. Semantic segmentation based on thresholding, morphological operation and using UNET model combine and developed to detecting the weed within and image Bottleneck for employing semantic segmentation is unavailability of labelled agriculture images at pixel level the system shows an effective and reliable classification of images captured by a camera. The image segmentation algorithm is very useful method in the image processing and it is very helpful for the subsequent  
processing. When the plants are separated from each other  
in the images, the results have been shown to be better. Also  
the lighting conditions are important to be able to make a reliable analysis.

1. **Reference**

<https://theailearner.com/tag/morphological-gradient-opencv/>

https://deepai.org/machine-learning-glossary-and-terms/feature-extraction

<https://www.cse.unr.edu/~bebis/CS791E/Notes/Thresholding.pdf>

https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5

<https://www.datasciencecentral.com/profiles/blogs/image-classification-with-hsv-color-model-processing>

<https://www.masterclass.com/articles/how-to-use-hsv-color-model-in-photography#how-to-adjust-hue-saturation-and-value>

<https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>

<https://www.learnopencv.com/histogram-of-oriented-gradients/>

<https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm>

https://www.mathworks.com/help/images/morphological-dilation-and-erosion.html