Detecting Weed Patches in Crops with Normalized Vegetation Index sensor using Support Vector Machines

Guilherme Manuel Oom de Albuquerque d'Orey

Student number: 201501385

March 2017

Master Thesis

Aarhus University

Department of Mechanical Engineering

Summary

The use of excessive herbicide has a negative impact on human health, environment and the competitiveness of agriculture. With the autonomous weed detecting systems relying on complex and expensive solutions, the purpose of this project is to study the possibility of using a normalized vegetation index sensor, that has already commercial use for the application of fertilizer, as a weed detecting system, as well as weed control on areas having weeds as only vegetation (pavements, urban areas etc,...).The possible features to train a support vector machine were tested, and the max accuracy reached was 78%, with 20% of the weed patches samples not being detected. It could also predict with near 70% of accuracy between high and low crop, with and without weeds. When the crop level was defined, the accuracy reached was of 80 % with only 10% of the samples of weed patches not being predicted.

This present study show that a robust sensor is capable of detecting weeds between crop rows until some extent, with a much simpler system than the existing ones.

Acronyms

ANN Artificial Neural Networks

Dif Difference between two consecutive NDVI values

FSA Feature Selection Algorithm

KS Kernel Scale

HM Harmonic Mean

ML Machine Learning

MM Moving Mean

MMED Moving Median

MSD Moving standard Deviation

NDVI Normalized Vegetation Index

NIR Near Infra Ref

RGB Red Green Blue

SVM Support Vector Machine

VIS Visible Radiation

List of Figures

List of Tables

Preface

This project is a thesis for a Master's in Mechanical Engineering, at Aarhus University. It was carried out from November 2016 to March 2017.

This project was carried out in colaboration with the Danish AgriFish Agency funded research project 'ØKOSPOR', work package 3. It balances multidisciplinary academic research and use of machine learning theories with a practical case study.

The project is worth 30 points on the ECTS scale.

Aarhus, 20 of March, 2017

Guilherme Manuel Oom de Albuquerque d'Orey

Contents

Chapter 1

Introduction

1.1 Background

Weeds are unwanted vegetation in the crop fields, due to its affect on the growing of field, uncontrolled weeds commonly reduce crop yields from 10 to 95% Press (2017),which makes weeding or hoeing a requirement for productivity and competitiveness. Weed control can be done with chemicals, by spraying herbicides, or mechanically, by cutting the weeds or hoeing.

The spraying of herbicides can be done all over the field, where is very inefficient when the weeds are not spread over all the field or can be done only where the are weeds, called as post emergence patch spraying (spraying herbicides only in those field areas where weed density or weed cover is above a given threshold) San et al., this has a negative impact, as the weeds start developing resistances or start to be tolerant to the herbicide, and has adverse impacts on both environment and human health Ahmed et al. (2012). The use of herbicides in Denmark, in 2009 was of 2807 tons bek (2009), having a serious impact on the cost, competitiveness, health and environmental consequences.

In hoeing, it's only done where there are weeds, has it tilts the soil, making the possibility of even more weeds growing(as some seeds that were underground come to the surface, receiving the sunlight needed to germinate) and can also damage the crop plants. So for hoeing to be efficient, a precise weed removing tool is necessary, which requires a weed identification method Nørremark(2010).

According to Nørr emark(2010), weed control compromise three key elements:

- A weed sensing system - identifying, localizing and measuring crop and weed parameters.
- A weed management model - applying knowledge and information about crop-weed competition, population dynamics, biological efficacious of control methods and decision making algorithms, and optimizing treatments according to the density and composition of weed species, economic goals and environmental constraints.
- A precision weed control implement, e.g. intra-row hoe or sprayer with individual controllable boom sections or a series of controllable nozzles that enable spatially variable application of herbicides.

This study is going to focus on the first element, as a lot of studies have been made(see section.1.1.2), but until now none has sufficiently efficient for commercial use, reason is that, sens-

ing systems cannot cope with natural variations of spectral or morphological characteristics and mutual shading between weeds. So site-specific weed control has major advantages, and there are many systems that are capable of precise weed control implement, but to be automated and efficient its needed a reliable weed detecting system.

(a)   (b)

Figure 1.1: On the left a tractor spraying a crop field, individual controllable boom section could be used to just spray the row with weeds, on the right a tractor with a hoeing device, this can be pushed up or down by a control system. With a weed detecting system can be programed to go down only when there are weeds autonomously

1.1.1 Literature Survey

1.1.2 Detecting weeds in Crops

A weed detection system integrates target detection sensors, data processing, and decision making systems, for the target detecting sensor, the ones usually used are detecting hyper spectral sensors, image sensors, spectrometers, remote sensing devices, thermographs, and laser sensors, Hong et al. (2012). The data processing usually uses as features shapes in the images, color indices, vegetation indices, or combinations of the three.

There are two main trends in computer vision to detect weeds, the first one tries to identify each plant as part of the crop or weed, the second tries to identify the pattern in the crop row and in weed spots De Rainville et al. (2012).

In Swain et al. (2011) uses the concept of 'active shape modelling' to identify weed and crop plants based on their morphology, achieving a 90% of accuracy. Active shape modeling is an algorithm where the shape of the leaf that is to detect is given, and a number of variations of

the shape are given to train the model. A segmentation, in the case a threshold from the excess green index, is done on the images to isolate the leaf shape from the soil. Even though the accuracy is high, this kind of system is very specific, for a particular kind of situation, having this accuracy for an ideal situation, where the weed species is known and is in a specific stage of growth. Ahmed et al. (2012) uses the support vector machine(SVM) a machine learning algorithm, to classify weeds and crops. Fourteen features from the images were tested to find the highest classifications rate, which nine were selected: solidity(are divided by the convex area), elongatedness(area divided by the thickness), mean value and standard deviation of the red component of the RGB, mean value and standard deviation of the blue component, and three moment invariants(for a more detailed explanation on the features, see Ahmed et al. (2012)). Ahmed et al. (2012) achieved 97% accuracy, with no miss classification(i.e the SVM got 97% of the classification right, and it didn't classified wrongly).Although had an high accuracy, SVM demands a high processing power to be trained, so it would only be practical to use in a very permanent situation, where the weeds are of the same species and size, otherwise, it would have to be trained regularly.

With the second trend, where the crop geometry is used, in Guerrero et al. (2012) the detection of the weed row is done without segmentation(i.e no filter is done on the image, to augment the contrast between the vegetation and the background or to isolate the vegetation), to serve as an alternative, since the segmentation is considered to be very demanding for the image processing software. Instead is done by dividing the grey scale image into various horizontal lines, and getting the points where is highest, and then with the Hough transform, to connect these points into lines, which then are considered the crop rows. The algorithm was successful the more rows the camera covered, and had some troubles when there were weed patches. In Tellaeche et al. (2011) a segmentation of the images is done, a binarization(i.e the pixels in the image, assume only two values, black or white). With the soil separated from the vegetation, the Hough transform is applied detecting the crop rows. Then separated in cells, where feature were extracted, as the area ratio of white in the left and right side of the cell, and the areas of white that are isolated from both sides. This features are processed and then used to train an SVM. This system was able to classify the patches to be sprayed with an accuracy between 60 to 86%. In De Rainville et al. (2012) a naive Bayesian classifier is used to discriminate crops from

weeds based on statistics computed from row and inter-row leaf area, where it detects the rows using the hough transform, and then applies an unsupervised learning algorithm, the Gaussian Mixture Model, to get features from the weeds. This system has two advantages, there is no need to train the algorithm and no prior knowledge of the crops and the weed species is required.This system had and accuracy of 90 to 95 % for various crops levels(the crops studied were corn and soy) and 70 to 90 % in classifying weeds. This system also requires a lot of processing power, making it expensive and slow.

The previous studies all rely on computer vision, in Ali et al. (2014) uses chlorophyll fluorescence induction curves with a Neural Network, to classify between weeds and crop rows, with an accuracy between 86 to 96.1%. Fluorescence induction means a series of changes in chlorophyll a fluorescence yield, observed in plant leaves when light is abruptly turned on after a dark period Ali et al. (2014). This has the disadvantage that has to be a video recording, since it's the changes that matter.

1.1.3 NDVI

Normalized Difference Vegetation Index is a index from the scale of -1 to 1, that measures the amount of living vegetation, but the practical values are from 0 to 1, so no vegetation is near 0 and 1 totally covered in live vegetation.
The NDVI is calculated from two measurements, the near infrared radiation,NIR, and the visible radiation,VIS, in the red spectrum, and then applying the formula Weier et al. (2006):

$$NDVI = \frac{NIR - VIS}{NIR + VIS} \tag{1.1}$$

Machine learning working with NDVI and Single Sensor Data

A research where SVM has been use with NDVI has been done. As NDVI is one of the indices in the LandStat NASA satellite system. In Zheng et al. (2015) uses SVM to classify crop species in Central Arizona, in the United States, using the NDVI Landstat time series. The features for the SVM were just the variation of NDVI along the year and location. Even though it's NDVI, this type of study resembles very little the problem for this Thesis, as it's a completely different scale, and the features for the SVM are different, as in Zheng et al. (2015) there are geographical

and weekly variations of NDVI, in the present study it's local variation over a very small area and variations in deciseconds.

Another particularity of the present study is that relies on a single sensor, so it's required to extract in the maximum features to train the SVM. In Ruiz-gonzalez et al. (2014) a SVM get the data from one uni axial accelerometer, to detect anomalies in an industrial harvester. Common problems simulated to verify if the SVM could detect them. The data from the accelerometer was processed into 12 features. Of these features, there was the mean value, median value, standard deviation, skewness, Crest Factor,...,(for details, read Ruiz-gonzalez et al. (2014)) and then used Exhaustive Search method to select the best feature subset. In Terzic et al. (2010) a single ultrasonic sensor is used to measure the fuel in a fuel tank of a car. The features extracted from the single sensor signal are Moving Mean, Moving Median, and Wavelet filter, but no feature selection was done. In Rajab et al. (2016) studies the use of a single-element piezoelectric sensor on a traffic lane to identify the type of vehicles. For the feature selection it uses the number of pulses, the ratio between pulse length and duration, and the principal component analysis.

So even thought there is no literature using a local NDVI sensor using SVM, it's possible to assume that the feature extraction used in the refereed studies can be applied to extract the maximum information from the single NDVI sensor, instead of just using the raw data.

1.2   Support Vector Machine

1.2.1   Support Vector Machine in comparison with other Machine Learning Algorithms

There are many classification algorithms, there are a lot of parameters that support the choice of each algorithms, such as the size of the data, the size of the labeled data, the amount of features, sparsity, class balance (if classes have similar amount of samples) precision required, processing power and time.

Even though, the dimension of the data and choosing the classifier is subject to a lot of discussion, many studies have shown that neural networks, decision trees and support vector machines (SVM) may often be able to classify a data set to a higher accuracy than conventional

statistical classifiers Foody and Mathur (2004)

The SVM has the advantage of not necessarily requiring a big data set, since only the data points on the border(i.e the data points that are closest to the hyperplane that separates two classes) are used to define the support vectors. Where the number of the support define by how easily the classes can be separated, this gives the SVM a high capacity for generalization with relatively small numbers of training data points and also the absence of local minima, and the sparse representation of its solution Ahmed et al. (2012). A small explanation on the most popular classifying algorithms is necessary to understand the comparisons that were made to lead to the decision of choosing the SVM, only supervised learning algorithms are going to be presented. A deeper explanation of SVM is going to be given in the next chapter.

For learning algorithms, there are two models, the discriminative and the generative models. In the discriminative models the algorithms learn the conditional probability distribution, i.e the probability of Y given X. The most popular discriminative models are Support vector machines, Neural networks and random decision forests.

In generative models, new data is randomly generating data values, from the joint probability distribution, of this models, the most popular is the Naive Bayes classifier.

The Naive Bayes classifier is quite popular for text categorization, where it calculates the category of text with the frequency of the words in it. In De Rainville et al. (2012) a naive Bayesian classifier is used to discriminate crops from weeds based on statistics computed from row and inter-row leaf area. One the biggest advantages of the Bayesen classifier is that is trained very fast, and adding a data point is just updating the probabilities of the features. Another advantages is that is an interpretable method, since the probabilities are defined, and can be checked to see which features have more impact in separating the class, as well to verify if there are features that are independent of the class.

SVM on the contrary is a black-box method, it's hard, or not straight forward to understand how the classifier got into the results. The main disadvantage with the Naive Bayes classifier is that it doesn't take into account the combination of features, so it's always the independent probabilities, and not combined ones.

The Random Forest or Decision Tree is a classifier that divides the data in the best possible manner, and repeats the step, until it reaches endpoints, that are the predicted classes. Random

forest has the big advantage that is very easy to interpret, since it creates a tree, where each node is a decision that was made to reach a certain class. One of the disadvantages of the Random Forest is that it's a linear classifier,so if a feature is not linearly divisible, it won't be capable to predict the class properly.

Artificial Neural Networks is a algorithm that tries to imitate a neural network in a brain, where each neuron has a weight, that is adjusted while the neural network is being train. ANN has the advantage of being able to handle complex non linear functions, which can find dependencies between different inputs(where Naive Bayesen can't) and allow incremental learning(i.e it's not necessary to run thought all the data, to train every time a new labeled data is given). There are two disadvantages, one is that as the SVM, the ANN is a black-box method, so it's not straight forward, or not even possible to understand how the algorithm got its results. The other is that ANNs can suffer from multiple local minima, whilst the solution to an SVM is global and unique.Also one of reason that SVMs often outperform ANNs in practice is that they deal with a big problem in machine learning which is overfitting(i.e overfitting, is when the model becomes too complex and becomes over sensitive to small variances in the data), SVMs are less prone to overfitting than ANN.

So SVM fits the requirements for this thesis problem, without having to much cost in complexity processing power, but choosing the algorithm is just the first step to create a classifier, that follows by preparing the data do train the classifier, which is as much or even more important than choosing the classifier, as it says in Foody and Mathur (2004), "Of particular importance are factors connected with the training stage of the classification. Indeed the nature of the training stage can have a larger impact on classification accuracy than the classification technique used".

1.2.2 Explanation of Support Vector Machine

SVM is a learning model that creates a hyperplane (which is a subspace with one dimension less than the ambient space), and this hyper plane separates two classes (that are going to be called positive and negative) with the biggest margin possible.
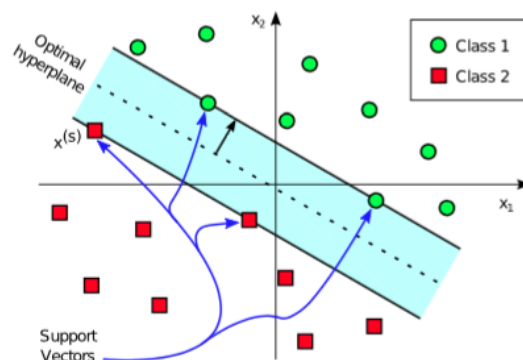

Figure 1.3: Visual Example of SVM with t

In fig.1.2 [1] it's easy to see the two classes, and the

hyper plane that separates them, the filled samples

are the support vectors, that located on the margin

and define the hyper plane. The Matlab version

used for this study has the Support Vector Machine

functions implemented, so to train the SVM is only calling the functions and input the classified

data, and the parameters.

The mathematical procedure, is from MIT OpenCourseWare (2014) and it's going to be demonstrated in the next section, for a better understanding of the functions and the working of the

classifier, to get the best use of it.

Mathematical explanation

Being the training data points given, $(\vec{x}_1, y_1), (\vec{x}_2, y_2), ..., (\vec{x}_n, y_n)$, where y is either equal to 1 for

the positive class and -1 for the negative class, to indicate to which class the point belongs.
The hyper plane equation is:

$$\vec{w} \cdot \vec{x} - b = 0 \tag{1.2}$$

Where $\vec{w}$ is a normal vector to the hyper plane, and $\frac{b}{w}$ is the displacement of the vector from the

origin(in fig.1.2 the hyperplane is the dashed line).This lead to the decision rule:

$$\vec{w} \cdot \vec{x}_+ + b \geq 1 \tag{1.3}$$

where x_+ belongs to the positive class

$$\vec{w} \cdot \vec{x}_- + b \leq -1 \tag{1.4}$$

and x_- belongs to the negative class.

If this equations are multiplied by y(1 for + and -1 for -), the result is:

$$y(\vec{w} \cdot \vec{x}_+ + b) \geq 1 \tag{1.5}$$

---

[1]Both picture 1.2,1.4,1.3 are from Ruiz-gonzalez et al. (2014)

$$y(\vec{w} \cdot \vec{x}_- + b) \geq 1 \tag{1.6}$$

Which are equal, so only one equation, that can be written as:

$$y(\vec{w} \cdot \vec{x} + b) - 1 \geq 0 \tag{1.7}$$

.The distance between the two margins is:

$$(\vec{x}_+ - \vec{x}_-)\frac{\vec{w}}{||w||} \tag{1.8}$$

If x is replaced with equations 1.5 and 1.6,

$$\frac{1 - b + 1 + b}{||w||} = \frac{2}{||w||} \tag{1.9}$$

As it was said before, the objective is to get the biggest margin possible, and that happens by minimizing $||w||$. Which is the same as minimizing $\frac{||\vec{w}||^2}{2}$ (which is more convenient mathematically, for later procedures)

To minimize $\frac{w^2}{2}$, which has the constraints(the equation 1.7), the Lagrange multiplier can be used, which lead to the following equation:

$$L = \frac{||\vec{w}||^2}{2} - \sum \alpha_i \big[y_i(\vec{w} \cdot \vec{x}_i + b) - 1 \big] \tag{1.10}$$

Taking the derivative of the equation:

$$\frac{\partial L}{\partial w} = 0 \quad (=) \quad \vec{w} - \sum \alpha_i y_i \vec{x}_i = 0 \quad \vec{w} = \sum \alpha_i y_i \vec{x}_i \tag{1.11}$$

$$\frac{\partial L}{\partial b} = -\sum \alpha_i y_i = 0 \tag{1.12}$$

Applying 1.11 and 1.12 in 1.10:

$$L = \frac{1}{2}\Big(\sum \alpha_i y_i \vec{x}_i\Big)\Big(\sum \alpha_j y_j \vec{x}_j\Big) - \Big(\sum \alpha_i y_i \vec{x}_i\Big) \cdot \Big(\sum \alpha_j y_j \vec{x}_j\Big) - b\sum \alpha_i y_i + \sum \alpha_i \tag{1.13}$$

$$L = \sum \alpha_i \frac{1}{2} \sum \sum \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \tag{1.14}$$

Solving this equation, w and b are known, which define the hyperplane.

It is important to notice that the function only depends on the vectors, which will be the support vectors.

Kernel Trick
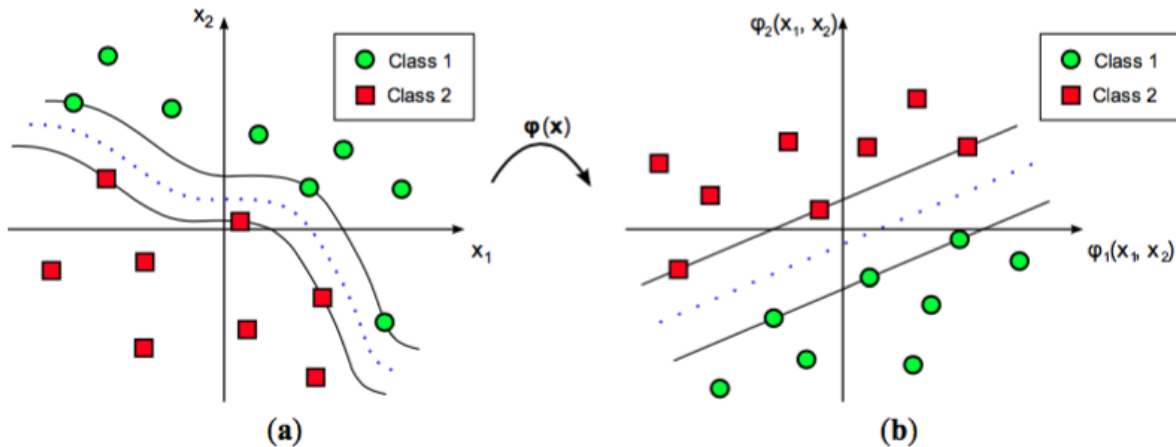
Figure 1.3: Visual example of Kernel function

As was demonstrated the hyper plane depended on the $\vec{x}_i \cdot \vec{x}_j$, but what happens when the classes are not separable by a liner hyper plane? A feature in the SVM is the Kernel trick, which give a lot of flexibility to SVM, it takes away the limitation of only linear separable classification.The Kernel trick is changing the coordinate system of the space vector $\vec{x}_i$ a function that can be called $\Phi(\vec{x}_i)$) is applied, the same for $\Phi(\vec{x}_j)$), a visual example can be seen in fig.1.3a and fig.**??**(b), is fig.1.3(a) is clear that the classes are not separable by a liner plane, so a new space is define with the kernel function, this is usually called the feature space . The kernel function is defined as:

$$K\left(\vec{x}_i, \vec{x}_j\right) = \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j) \tag{1.15}$$

There are various Kernel functions. In the table.1.1 are the other 3 most common kernels:

These are the kernel functions used in this study.The polynomial used is of second order(n=2) and third order(n=3).

Kernel Name	Function
Gaussian or Radial Basis Function	$exp\left(2||\vec{x}_1 - \vec{x}_2||^2\right)$
Linear	$\vec{w} \cdot \vec{x}_i$
Polynomial	$\left(\vec{w} \cdot \vec{x}_i\right)^n$

Table 1.1: Kernel functions
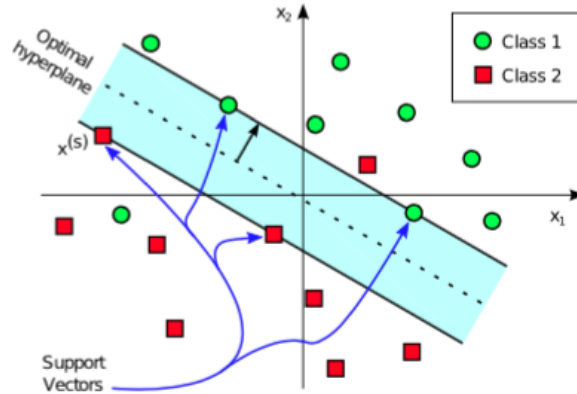
Figure 1.4: SVM with inseparable classes

Hard Margin and Soft Margin

In the previous sections, it was assumed that the classes can be completely separable, this is what is called the hard margin(i.e where all the points belonging to a class are respectively on the side belonging to its class, fig.1.2 is an ilustration where hard margin is applied). This is the optimal situations, but is not always possible, in that case a soft margin is used(1.4 is an illustration where soft margin is necessary).

In eq.1.10 the objective is to minimize $\frac{||\vec{w}||^2}{2}$, a variable ξ is added to eq.1.7:

$$y(\vec{w} \cdot \vec{x} + b) - 1 + \xi \geq 0 \tag{1.16}$$

ξ is the tolerance of the margin. But if ξ is very large, the eq.1.16 would always be satisfied. So a constant C is added to penalize if ξ is very large. Applying in eq.1.10:

$$\frac{||w||^2}{2} + C\sum\xi \tag{1.17}$$

So C is controlling the amount of points that are in the wrong side of the hyperplane. If C is very

low, a lot of support vectors are defined, which mean the algorithm is not generalizing(i.e if a big ratio of the points are support vectors, it means most of the points are contributing to the decision). If C is too big then it might over fit(when the model becomes too complex and has big variations, and becomes to sensible to noise) . To find the optimal balance:

$$min\Big[\frac{||w||^2}{2} + C\sum \xi\Big] \tag{1.18}$$

So the first term is maximizing the margin between the classes, the second part penalizes the points that are on the wrong side of the classes, and C is balancing the two parts Foody and Mathur (2004).

Multiple Class SVM

In the explanation there was only two classes, and SVM is a binary binary classifier, but it can be applied into multiple class problem by dividing into multiple binary classification problems. There are two ways to do this, "one against one" or "one against all", being "one against one" where the classes are paired and there is a classifier for every pair of classes, and the class that has more wins(i.e predicted more times) is the one selected. In the "one against all" where the classes are distinguish from all the rest of the classes, where the classifier predicts is the sample belongs to a class or not and the strongest prediction is selected. In this study, for the multiple class, the one against one is used.

It's a binary class method,but can be applied for more than two classes, by using "one against one" or "one against all", which are ways to decompose a multiple class to various sub binary class, being "one against one" where the classes are paired, and the "one against all" where the classes are distinguish from all the rest of the classes.

Feature selection

Feature selection is choosing the subset of features that better represents the data and produces the best classification. It's a crucial step on a classification process, as it reduces the processing requirements, reduces the noisy features(i.e features that are independent of the classes), maximizes the performance of the algorithms and it simplifies the use and update of the algo-

rithms Salappa et al. (2017). In Salappa et al. (2017) performs an analysis of feature selection algorithms(FSA). The FSA, have various ways to select the relevant features, based on in class dispersion and between class dispersion, and selects the features with four processes, in forward selection it starts with no features selected and adds the features while it improves. In backwards elimination it starts with all features and eliminates while it improves, in compound it's a combination of both, and then there is random selection. Then it compares various FSA, and then compares the accuracy of the algorithms with the selected features to the complete features set. With SVM the accuracy got on average 6% higher the selected features than with the full feature set.

1.3   Objectives

As analysed in the previous sections, the weed detection systems that have been studied, are only efficient in ideal and specific conditions and required specific systems and equipment, solely for the use of the weed detecting system, requiring an initial investment.

In this Master Thesis, it's going to be studied the possibility of using a normalized difference vegetation index(NDVI) sensor to detect weed patches.

NDVI sensors are common in farming for controlling the application of fertilizer Samborski et al. (2015), so it's a sensor that is commonly use, and it's robust and accessible. Making these sensors capable to detect weed patches would create the opportunity to have a automated weed detecting system that is simpler and robust than the ones that are currently available, allowing a big reduction on the herbicide usage, making agriculture more economic, sustainable, efficient and ecological.

The main objectives of this Master's Thesis are:

1. Verify the possibility of detecting weeds with a NDVI sensor.
2. Study the best features of NDVI data to detect weeds.

1.3.1   Limitations

This a Master Thesis, so there are limitations. The time limitation is quite considerable, as the it should be done within a semester. As this is an experimental work, one semester the number of testing is limited, and since it depends on the crop grow, the number of possible repetitions becomes one. Passing with the measuring instruments in the crops can not be done over and over, as the wheel move some some soil over, after some passages with the quad bike some plants get under the soil. Its also required some specific weather conditions, has if it's raining passing one the crop could get the plants cover in mud.

This report had to be handed-in in March, and it happen that the crops only grow enough to be measured in the end of January, which limited the time for the data treatment and processing and application of the algorithms. In matter of resources, the instruments used were the ones available, which were not the state of the art. Another limitation on the measurements was the power supply of the measuring instruments, which were powered bu the battery of the

quad bike and there were some moments were the measurements were interrupted due to the instruments, not receiving enough power.

1.4   Structure of the Report

The report is structured as follow. In Chapter 2, the experimental set up is explained, how the sensors were mounted and its technical specifications. A description of the field is also given.

In chapter 3 the the data processing methods are explained and the classification and feature selection is done and explained.

In chapter 4 the results of the trained SVM are displayed in tables and confusion matrices.

In chapter 5 the discussion of the results is done and proposal for deeper research are made. Then there are appendix to for more detailed explanations, with the bibliografy at the end.
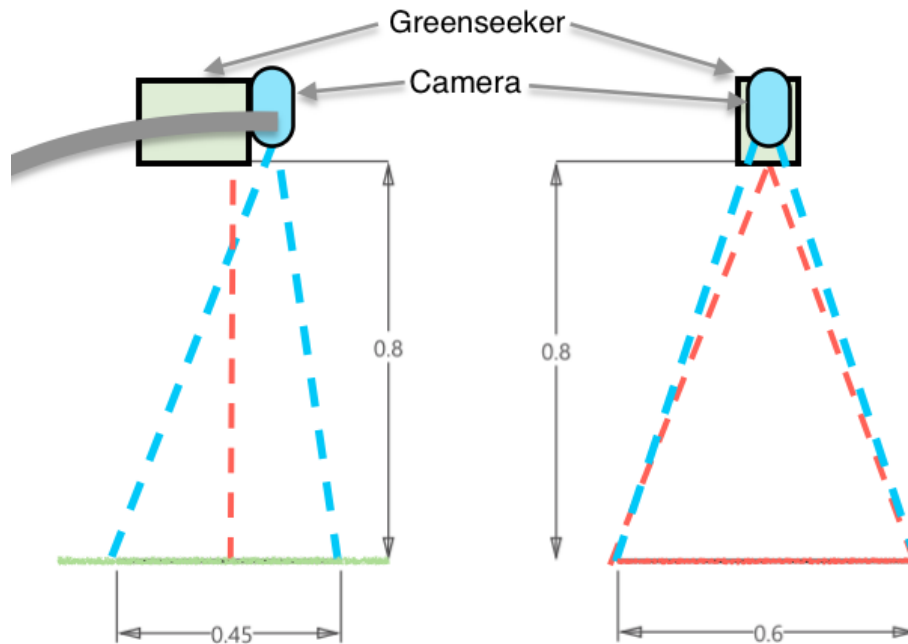
Chapter 2

Experimental Set Up

Figure 2.1: Diagram of NDVI sensor and camera

The data collection was made in the Research Centre Follum, in Viborg, of Aarhus Universitet, on the 26 of January of 2017. The measurements were done in a experimental cultivation, of winter wheat, and the weeds, of two leaf weeds(a variaty of species), were planted on specific area, for the purpose of the experiment. A layout of the crop field can be seen in fig.2.2c The inter row width is 0.2 meters. The five spots were defined to have a big variety of measurements, spots where the crop was more develop with and without weeds, low crop development with and without weeds, spots where the crop lines were with a lot of interruptions, where crops plants growing in between the crop lines. On the end of every spot of the measurement a white board placed to serve as reference in the measurements (since the NDVI goes to zero when reaches the white board).

The NDVI sensor used was the GreenSeeker® RT100, by Trimble, was doing 10 measurements per second, within a line of 60 cm wide, perpendicular to the crop row(can be seen in fig.2.2a A metal frame was fixed to the quadbike, high enough for the camera to catch the whole line of measurements, 0.8 m from the ground so it catches the whole width of the NDVI measurement, this was verified with the red light emitted by the GreenSeekerl, a picture of it can

be seen in in apendix A.2,the GreenSeeker needs to be between 0.8 to 1.22 m, according to the manual). The camera used was a Marlin F-033, by 2.1, it recorded color images with the size of 640x460 pixels. The camere lens covered a 45 cm by 60 cm are on the soil surface. The software that received the images was AVT Firepackage Smart View V2.3.1.

Both the GreenSeeker and the camera were directly connected to a laptop, that recorded the data. The set up was installed in a Quad bike, which supplied the power from it's battery.In 2.2(a)(b) a picture of the quad bike with the GreenSeeker, the camera and laptop. The Quadbike was driving approximately $0,8m/s$.

Every plot was measured 3 times both ways.

All data processing was performed offline using the academic software package of MATLAB 2016a. The Data processing will be treated in the next chapter.

Camera	Marlin F033
Sensor:	Sony ICX414, progressive scan CCD progressive scan
Resolution:	656x949
Max f.p.s:	73

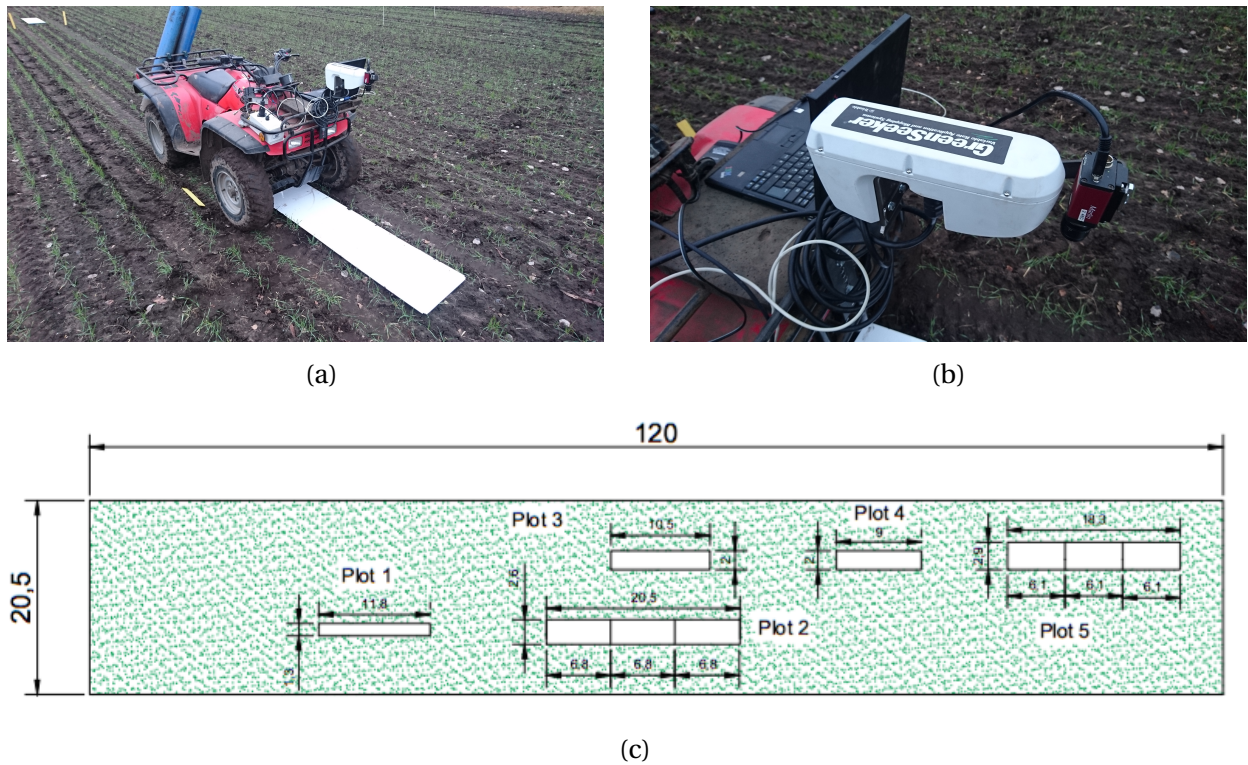Table 2.1: Camera Specifications

(a)

(b)

(c)

Figure 2.2: Set up of the measurements, the camera and the NDVI sensor can be seen, with the white board right in front of the Quad bike in 2.2a. In 2.2b a focus on the GreenSeeker(white sensor) and the Camera(in red), and the laptop behind. In 2.2c the layout of the experiment field where the measurements were done.

Chapter 3

Processing the Data

3.1 Data Processing

In this chapter, the classification, the processing of the data and the feature selection is going to be explained.

3.1.1 Classification

Three classifications schemes were given to the pictures, even thought the objective is just to detect weed patches, there might be approaches were the classifier is capable of separating the data into sub sets, and then divides the subset again, this is what is called as hierarchical learning. It gives the advantage that it might be easier for the classifier to separate the data into two sub sets and then separate it again than separating all at the same time or process, and it has more flexibility since the features used for the first classification can be different from the features that separate the second, and so on, for this case the classification scheme is going to be referred as HC for classification between high and low crops, and second level to classify between crop or weed patch, for both high and low crop.

The three classification schemes, are listed in 3.1.1.

- Presence of weeds or not(BC)
    - Crop (c)
    - Crop with weed patch (c wp)
- Presence of weeds considering the Crop Density(MC)
    - Low crop (lc)
    - Low crop with weed patch (lc wp)
    - High crop (hc)
    - High crop with weed patch (hc wp)
- Hierarchical learning (HC)
    - Low crop (lc)
        * Crop (c)
        * Crop with weed patch (c wp)
    - High crop (hc)
        * Crop (c)

∗ Crop with weed patch (c wp)

The classification of the images was done manually, individually to each image(for more information see appendix B.2).In table 3.1, the amount of classified images to for each class. While on the classification depending on just if there were weeds or not, was much more straight forward. The classification used in MC and HC is the same, just separated in different classes, while an image that was classified in MC and HC was also included to the corresponded class in BC, the opposite is not true, that explains the difference in the number of classified images for each scheme(for ex. the crop could be very dense in a line, but it was missing or was not small in the other lines. Examples of images from the classes can be seen in fig.3.1.
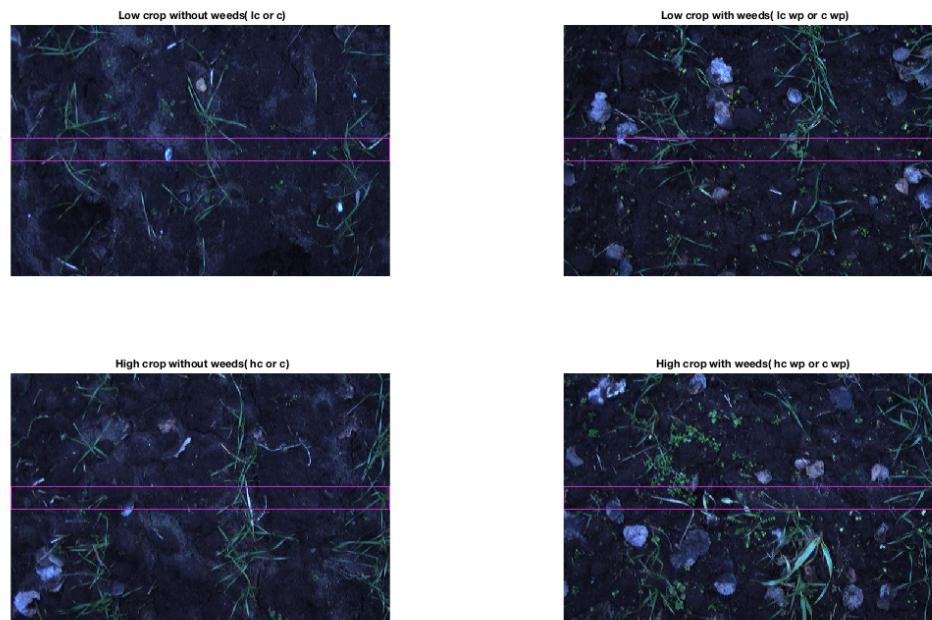
Figure 3.1: Images from the the camera of each classes, the rectangle in the middle was added for reference where the NDVI was measuring

3.1.2 Data Synchronization

The data from the NDVI was stored in the laptop hard-drive in a comma separated value file. The camera pictures were stored in JPG file foramt. The data from both sensor had to be synchronize, and since they were logging independently, a times tamp was put on every data log. . As said

Crop Spots	BC		MC				HC			
							LC		HC	
	C	C WP	LC	LC WP	HC	HC WP	C	C WP	C	C WP
Plot 1	199	9	88	9	37	0	88	9	37	0
Plot 2	252	411	110	181	93	225	110	181	93	225
Plot 3	114	38	34	33	60	0	34	33	60	0
Plot 4	15	4	13	0	02	2	13	0	02	2
Plot 5	575	28	151	15	194	9	151	15	194	9
Total per Class	962	481	395	237	286	236	396	237	286	236

Table 3.1: Number of samples of each class

in the previous chapter, there were some limitations(both power supply and saving the data on the computer, was not always constant) and that influenced the data. Also the time stamp varies with seconds, and even thought the NDVI and the Camera were recording 0.1 s and 0.75s respectively, the seconds fraction was added later. So the time stamp had to be manipulated for the synchronizations between the images and the NDVI measurements.

A synchronization is needed, as the classification was done on the images but the objective is to apply to the NDVI. To synchronize, a color index on the image was used, to verify its variation along the measurement and to compare with the variation of the NDVI. There were five color indexes used to study the variation along the measurements. The color indices's used were the sames used in Guerrero et al. (2012) and another color index was used, that was made particularly for this study. The color index done for this study in particular case was the one that showed the best correlation, which can be seen in figure 3.2, with both indexes normalized. The top plot is the color index for the whole image (the sum of all the pixel values) and the bottom one is only the rectangle where the NDVI is being measured(the red stripe in fig.2.2b), as can be seen there are parts where it each one correlates better with the NDVI. A more detailed explanation on the color indexes can be read on appendix A.2.1.

In fig.3.2 to compensate for the difference in the frequency of the NDVI and images, the classification given to an image, was given to the NDVI measured $\pm 0.1s$ of the image (i.e for every 3 images taken, 4 NDVI values were measured, so the classification of the ith image was given approximately to the $i^{th \pm 1}$ NDVI measurement, since the classification was done on the images, and the images overlap).

In fig.3.3 and .3.4 can be compared the two way to pass the classification to the NDVI, in
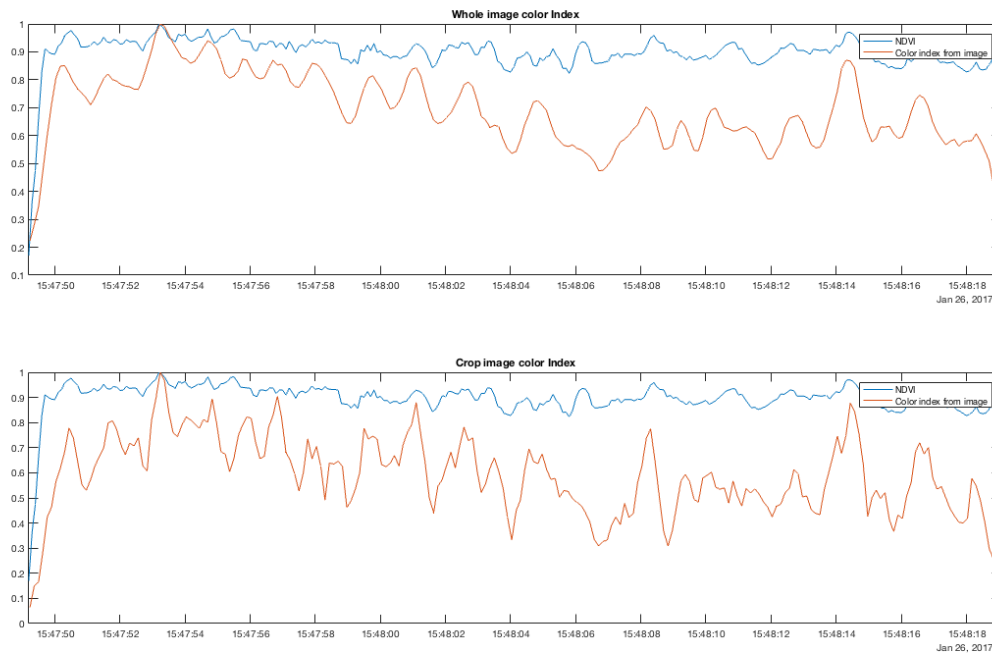
Figure 3.2: The top plot is from the whole image, the bottom plot is only the area where the NDVI is being measured

this case the time stamp is continuous, which is not always the case, in those cases the data was excluded(for more information see appendix B.2. Is also important to note that differences between the plots gets bigger on the left side of the plot, which show that the data log was not always constant in time.

One observation that can be made from fig.3.3 is that the Color index of classified images is higher for the Weed Patches Classes, which validates the classification of the images, this is not always true for the NDVI, but it also varies less.

Only the time synchronized classifications are going to be used.

3.1.3 Features

As seen in the previous chapter feature selection is a crucial part to train a classifier. By looking to fig.3.6, it's possible to see the NDVI values labeled, adn see some trends in each class, as the mean value, variation, pikes, which will be used as features, and using some from Ruiz-gonzalez et al. (2014). The list of features is in the end of the section. Some features depend on the neigh-
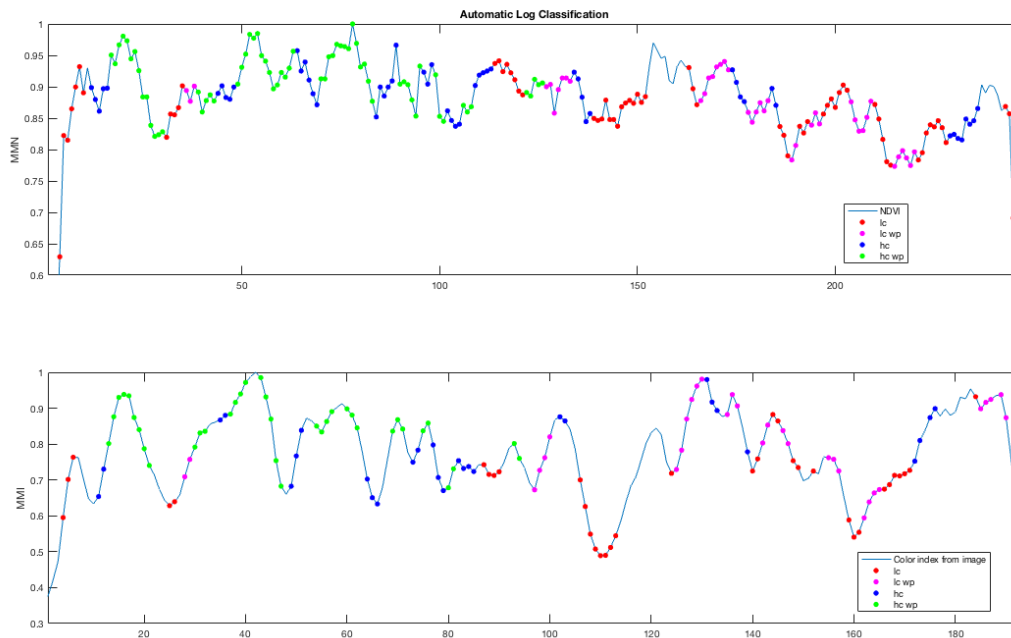
Figure 3.3: The data collected at Plot.5, in the top the NDVI classified by log order. In the bottom the color index.

bouring values, as for example the moving mean is the mean between n_{-i} to n_i, this interval is constant for the same set of features, the size of the interval is going to examined in the next section.

Figure 3.4: The same data from fig.3.3, but with the time stamp as the horizontal axis

Figure 3.5: Examples of the labeled values of NDVI

- NDVI value
    - See eq.1.1
- Moving Mean(MM)
    - $\frac{NDVI_{-i}+...+NDVI_i}{N}$
    - The sum of the NDVI values, divided by the N, the number of values
- Moving Harmonic Mean(MHM)
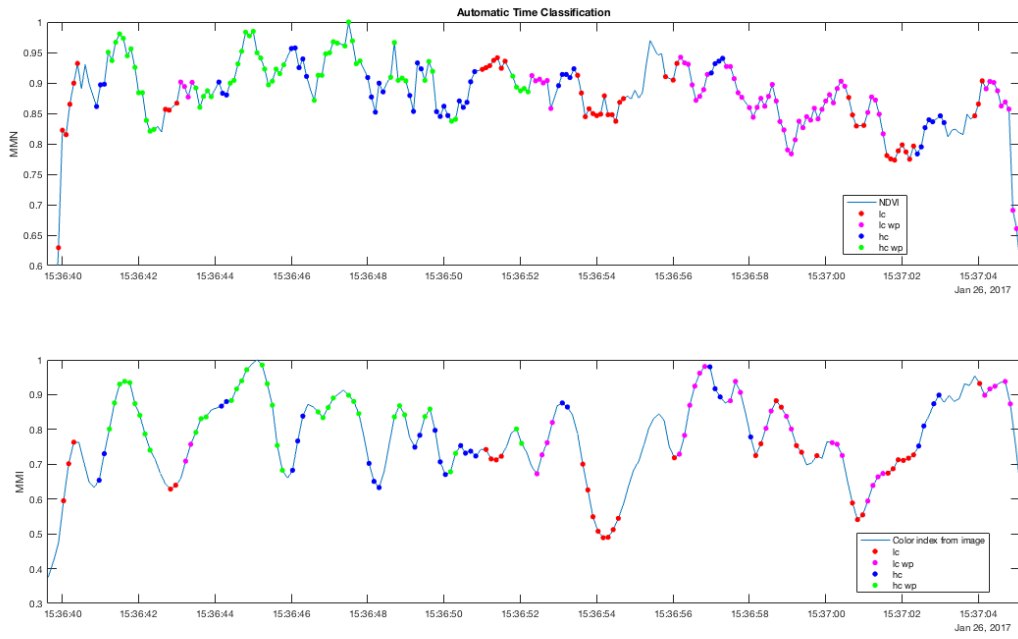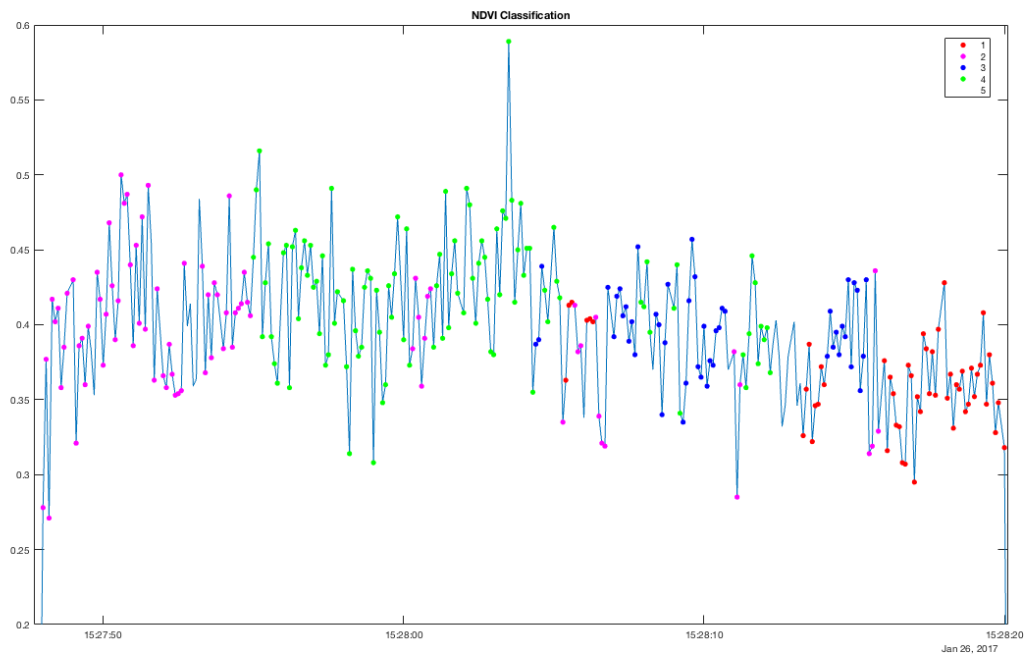   - $\frac{NIR-VIS}{\frac{1}{NDVI_{-i}}+...+\frac{1}{NDVI_i}}$
    - The number of values divided by the sum of the inverse of the NDVI values
- Moving Standard Deviation(MSD)
    - $\sqrt{\frac{1}{N-1}\sum|NDVI-mean|^2}$
    - The Standard Deviation is the square root of the variance.
- Range
    - max(NDVI)-min(NDVI)
    - the difference between the maximum and the minimum of the interval
- Moving Median(MMED)
    - The NDVI value that separates the interval in two parts.
    - The Standard Deviation is the square root of the variance.
- Difference(Dif)
    - $NDVI_i - NDVI_{i-1}$
    - The difference betweeen two consecutive values of NDVI

3.1.4 Feature selection

To verify the size of the interval that is used, the models, for the three schemes are run with full features, and the accuracy is calculated. Then the the interval starts with 3 samples, and the accuracy is calculated, then the one sample is added, and the accuracy is calculated again, this is repeated until the interval is composed of 30 meassuremnts. The accuracy is calculated by using a 10 fold cross validation(see appendix A.1). In fig.3.6 the normalized variation of accuracy for the three schemes is shown, and for the Multi classification the maximum accuracy is with 6 NDVI values, but has another pike with almost the same accuracy with 11 values. In Binary classification the maximum is with 12 and 13 values, in Hierarchical Classification is at 11. To standardize the interval, and since the differences are are very small, the size chosen is

Subsets of Features	NDVI	MM,MMED,Range,MSD*	MM,MSD	MMED,Range,Dif**	All
Binary Classification	74.36	78.19	79.81	77.94	77.76
Multi Classification	44.36	51.05	50.15	50.25	50.85
HC between hc-lc	59.58	61.69	59.58	63.79	60.88

Table 3.2: Accuracy percentages of the subsets that got the best results.*The subset with HM,MMED, Range, Dif got the same result in MC,** the subset MMED,Range,MST got the same result in HC

12 measurements for the three. To select the features the exhaustive method is used Salappa
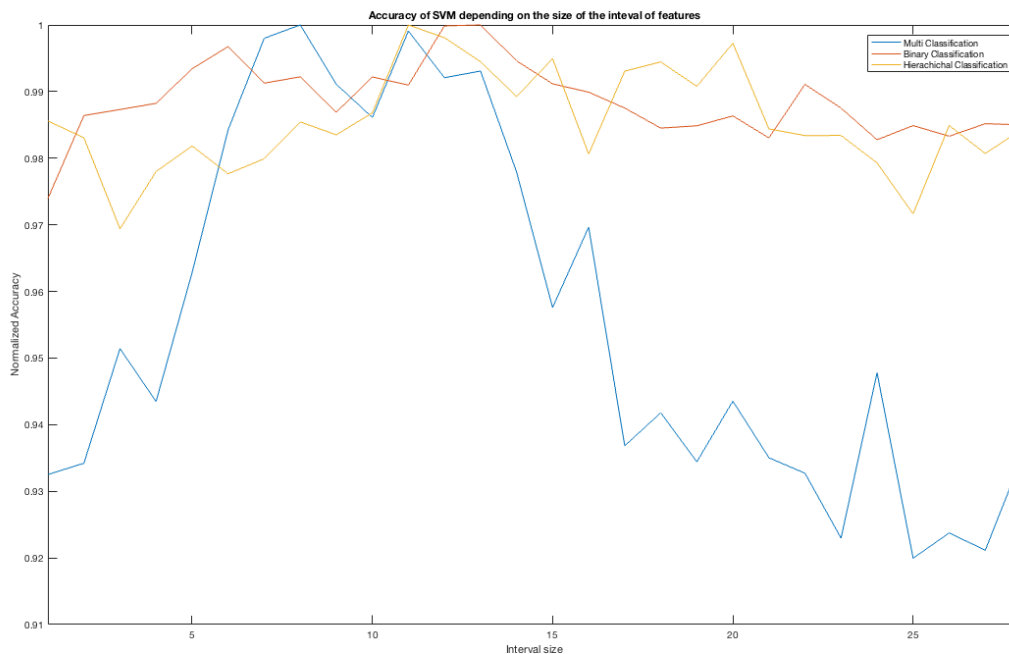
Figure 3.6: Variation of Normalized accuracy with the variation of the number of samples used to calculate the features

et al. (2017), with a 10 fold cross validation. In tab.3.2 the subsets of features that got the best result for each classification scheme are shown. The accuracy with only the NDVI and using all features is shown for comparison.

As can be seen, the values don't vary much, which can be explained since all the features are derived from a single one, the NDVI. It's also important to note that the percentages between the classification schemes can not be compared, since the binary classification has much more samples.

Sub Set of Features	NDVI	MM,Range,MST	NDVI,MM,MED,Dif	All
Low Crop	69.53	75.32	72,75	72.96
Dense Crop	77.21	79.85	84.18	81.73

Table 3.3: Accuracy percentage for the second level of the Hierarchical

Aproper comparison between the schemes is going to be made in the next chapter.

The features used are going to be the ones that got the best accuracy, for the Binary classification only the Moving Mean and Moving Standard Deviation, for the Multiple Classification the Moving Mean, the Moving Median the Range and Moving Standard Deviation. For the Hierarchical Classification the Moving Median, Range and Difference, for the classification between Dense Crop or Low Crop. The advantage of Hierarchical classification, is that different features can be used for different level of the hierarchy, and for different branches of the same level, in tab. 3.3 the

Chapter 4

Experiment Results

Kernels	Linear	2nd order polynomial	3rd order polynomial	Gaussian
Binary Class	0.5571	33.0701	0.1288	0.5958
Multi Class	1	1	1	1
HC	0.3600	2.3109	0.0222	1.9770
HC lc	0.0552	0.0254	0.3618	0.8602
HC hc	287.4476	910.0951	0.5289	0.0031

Table 4.1: SVM optimized C parameter

With the features all selected the SVM can be trained. As explained in the chapter before the Binary Classification has more samples, so in the first section of this chapter, the Binary classification will be trained with the same sample size as the others for comparison, in the section after the Binary Classification is going to be trained with all the samples available.

4.1 Parameter Values

With the features selected, it's possible to adjust the parameter values discussed in section.1.2.2. The parameter were found using random search, which is an numerical optimization method that selects a random point in the optimization space, and searches for the best(inn the optimization parameter) within a specified radius, and and keeps the best one, then searches again from the selected point, in this study this procedure was repeated 30 iterations. The best parameters, for C are: The very high value means that the data is separable, as seen in section **??**. The values that are equal to 1 is because the optimization for the mutli class SVM was required to much computational power, so the default value was used. The values that are low, it's because the data was not separable.

4.2 Comparison of all Classification schemes

In tab.4.2 the accuracies for the 4 kernels tested is shown. The kernel with more accuracy overall is the Gaussian Kernel. To distinguish the crop or a weed patch, when the crop level as been defined the SVM has the highest accuracy, with a Linear kernel being having 83% accuracy. In tab.4.4 the results are from the binary classification with all the samples, but the difference is

not very big, compared with reduced one, which agrees with whats was said in section 1.2.1, that SVM has a good generalization capacity.

Kernel Functions	Linear	2nd order Polynomial	3rd order Polynomial	Gaussian
Binary Classification	77.6	59.0	57.25	78.4
Multi Classification	53.4	56.2	67.8	67.8
Hierarchical Classification	63.2	62.6	49.5	65.9

Table 4.2: Accuracy, in percentage, with different kernels.

Kernel Functions	Linear	2nd order Polynomial	3rd order Polynomial	Gaussian
Low Crop	74.5	74.5	62.9	80.3
Dense Crop	83.6	82.1	81.7	80.4

Table 4.3: Accuracy, in percentage, for the second level hierarchy

Kernel Functions	Linear	2nd order polynomial	3rd order polynomial	Gaussian
Binary Class	79.6	76	58.7	79.8

Table 4.4: Accuracy, in percentage, of the binary class with all samples
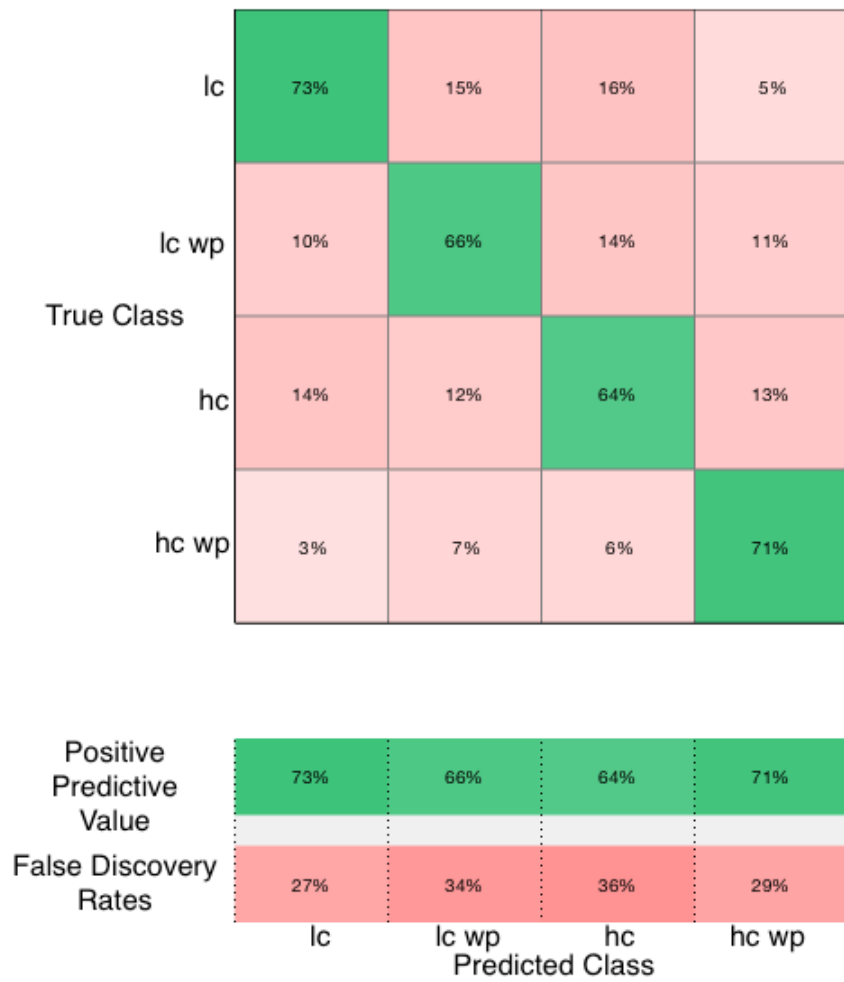
4.3 Confusion Matrices

The confusion matrix is a very simple and clear way to evaluate how good a model can classify accurately each class. A confusion matrix is constructed using cross validation, the columns are the values that are used to test and are distributed in the column of the class that the model predicted. The rows correspond to the True values. So the best classifiers have the values in the diagonal, that is, where the column of the predicted class meets the correspondent True class row,for example, comparing Binary Classification,fig.4.1a to Hierarchical Classification,fig.4.1b, it's possible to see that BC got higher values than HC in the diagonal,and has been seen in tab.4.2 that BC got better accuracy.

More information can be taken from the confusion matrix, for example, in this study is much more relevant that the SVM doesn't miss a weed patch than if it detects a weed patch when is just crop, since in the first it allows the weed to grow and on the second it cuts or sprays herbicide. So to look for the Classification scheme that is more appropriate for weed detecting, one should look for the columns of predicted crops without weeds, and for the row of true weed patch values.

(a) Binary Classification

(b) Hierarchical Classification

(c) Multi Classification

Figure 4.1: Confusion Matrices for BC,MC and HC

(a) Classification in Low crop

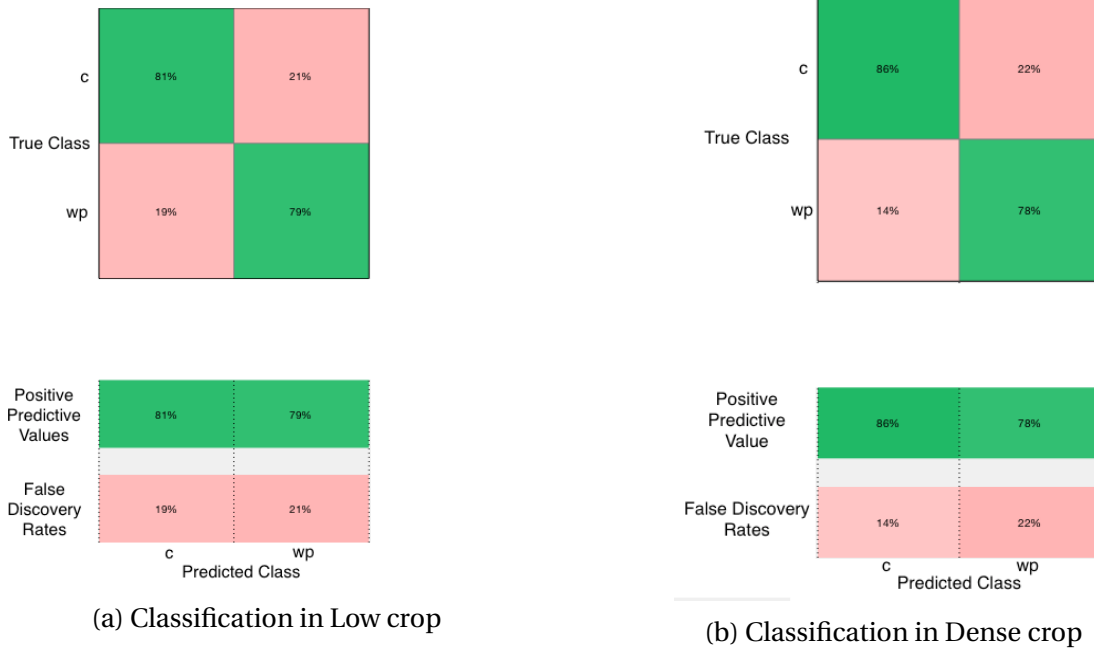(b) Classification in Dense crop

Figure 4.2: Confusion Matrices for second level of the hierarchy

Figure 4.3: Binary Class with SVM trained with all samples

Chapter 5

Summary and Recommendations for Further Work

5.1 Discussion

5.1.1 Features

All the classification schemes gave a different subset of features, that gives meaning to the classification schemes, even thought the accuracy with difference subsets did not vary much($\pm$1%), and since that accuracy is also an estimate(from cross validation) the feature selection gives the security that the SVM is being train with some consistent features, but also doesn't allow to get a better insight of what can define each class.It's also interesting to see that the moving mean is a feature present in all subsets, less the fist level of the hierarchical, which is to distinguish between high and low crop.

The difference between only using one feature (the NDVI) than use all features is also not very big(7% at the second level of the hierarchy with low level crop).

The subset to detect weeds on a low crop density (second level of the hierarchy) has the range, moving mean and moving standard deviation. This is might be explained, that with low crop, there might be measurements where the NDVI is only reading soil, so a very low value(since the crop rows are not constant, see fig. 3.1) and then can be reading crop and a weed patch, a high value, which might be a reason for the range to be a good feature for that classification.

The size of the interval that defines the features is also interesting,since all of the schemes got a similar one, since this has a relation with the velocity of the quad bike, it would be interesting to understand how much it would vary with the quad bike at a different speed .So it's possible to conclude that although the results wouldn't be substantially different if no feature selection was done, the FSA choose the best subset, improving to the classifiers.

5.1.2 Classification Schemes

Unfortunately the SVM is a black box, so it's not straight forward what "decisions" the SVM made, even thought a lot of information can be taken out of the results.

The classification scheme that got the best results was the Binary Classification. What was interesting is that the training with the whole samples only added 1.4% to the Gaussian Kernel, the big difference was for the 2nd order polynomial, that added 17%.

Since the difference between the samples was more than double of the size(1443 to 632), this quite agree with the fact that SVM is able to generalize over small data sets.

Th multiple classification had some poor results but being the classes in the middle that had lowest scores, by middle is that it can be assumed that the class with the lowest NDVI value is the low crop, and the highest is the high crop with weeds, so the other should be in the middle, and that's probably the reason for the low scores. It classified 24 % of the lc wp as without weeds, and 9 % of the hc wp as without. Being these values the most critical if this was to be a real weed detection. These classification scheme was purposed to see if if there was a trend in that isolates the weeds and the crop level, and there might be, because all the class, besides lc wp, had more miss classifications with the corresponding class of the other crop level , (i.e. true lc had more predicted values of hc than lc wp and hc wp, and true hc had more predicted values of lc than the ones with wp, and true hc wp had more predicted values of lc wp than lc or hc), which means the SVM is having more trouble in separating the the crop densities than the crop from the weed patches. This becomes more clear on the hierarchy classification. The Hierarchical Classification got some interesting results on the second level, when it was actually to detect the weed patches, being the classifiers with the highest accuracy.

But it got very poor results to distinguish between the crop density, but this again might be also due to the fact that the classification was done manually, so the values from the same class could fluctuate a lot. On the contrary for the second level of the hierarchy, it's straight forward, which also might explain the results. It's also the only classification where the Linear Kernel performed better than the Gaussian.

On the second level hierarchy its very interesting that such high values were achieved, which mean that with a more accurate reading of the crop levels there might be a clear trend in the values with weeds and without weeds, which is what this study is focused on.

Overall the accuracies achieved are consistent and that the algorithms were properly applied.

5.2   Conclusion

The result of this experiment given the conditions is quite positive. With a single sensor was possible to detect weeds with 78% accuracy, and with 20 % of not detecting the weeds (the binary classifier predicted it was only crop 20% of the weed patch samples), which is a promising number considering the advantages in comparison with the other solutions (mostly using computer vision, see chapter.1), in Guerrero et al. (2012) an acuracy of 93% was achieved, in Tellaeche et al. was between 60% to 85% and in De Rainville et al. (2012) between 85% to 95%, all with computer vision, and computacional demanding algorithms, like feature detections in images. The 78% of accuracy is not far away from those.

The Gaussian Kernel was by far the most accurate kernel function, this means that Linear classifiers would have been in disadvantage. Choosing the SVM with various kernels gave the possibility to verify that.

THe SVM was very capable of doing the task required, as was seen, in the Binary classification, the number of samples double, the accuracy didn't got much higher, only 1% more. From this, it's possible to assume that the other classification schemes were not very influenced by the number of samples.

But even though the number of samples was enough, the quality of them probably had more effect on the classification accuracies.Since the classification was done over periods of time and not specific samples, this might have added some noise to the data.

Over all, doing the feature selection, test the various kernels and various classification schemes proved to give new insights on the possibility of detecting weeds with just an NDVI sensor.

5.3   Recommendations for Further Work

The study for the NDVI as a weed detecting sensor has still a lot of possibilities, some ideas are worth to explore:

- Parallel measurements, collect data from two or more sensors simultaneously. For exmple two NDVI sensors are mounted on a beam in a tractor, measuring the crops, one of the sensors passes over a weed patch while the other not, would a ML algorithm be able to detect the difference in the measurements?

- A relation to the velocity of the vehicle (that carries the sensor), the features in this study depended on an interval of measurements, studying how the velocity that the sensor passes over the crops might would give a better understanding the effects a weed patch has on the measurements.

- Precise measuring, the biggest limitation on this study was the lack of precision between the images and the NDVI, which made the classification on the NDVI in intervals than in specific values, the fact that was not possible to know exactly what the NDVI was reading at the specific sample didn't allow to study deeply the difference between a value from a weed patch from a crop(i.e is the difference between defining that these samples from this interval are when it's passing over a weed than defining, values X to Y is of a reading of a weed), then would be interesting to study deeply the features from each samples.

5.4   Summing up

Overall SVM is was more than capable ML algorithm to process data from a sensor like the NDVI. It got a very reasonable results, not far from the computer vision ones, and as it's a algorithm that its trained offline, and it doesn't need to save the old data (although it doesn't implement new data to learn). It can be applied in a sensor like the GreenSeeker foe weed detection. Although some more research needs to be done to verify if higher accuracies can be reached.

Appendix A

A.1   Cross Validation

Cross Validation is a validation technique very common statistics, and used in machine learning to test the accuracy of a classifier. In Cross Validation, the data set is separated into subsets(usually called as folds) and and then a sub set is used to train the model, while the other is given to the model as new data, which then the model predicts, and then is verified how the model behaved with new data, since this data was actually already classified, the accuracy is the fraction of samples correctly classified.

The concept is simple, but there are considerations in how to divide the Data set, as it can choose a subset that doesn't represent the data very well and if the subset chosen to train the model is to small, the model might not be able to generalize.

Of Cross validation types, there are three that are the most common ones. The leave-p-out, where p observations of the data set are chose to validate and the rest is chosen to train the model. This is repeated until all samples have been part of the validation subset. There is another type, which is a particular case of leave-P-out, which is leave one out, where one sample is use to validate and the rest is used to train, and it's repeated by the number of samples.

The other type the K fold, where the data is divided k times, and only one is used as the validation, and is repeated k times, if k is equal to the number of samples, it's the sames as leave one out.

Accuracy

The accuracy A of a classifier is given by:

$$A = \frac{N_{correct}}{N_{total}} \tag{A.1}$$

In cross validation the accuracy is the average of the accuracy for each subset.

A.2   Color Index

A.2.1   Matlab Created Color Index

The color Index that was created for this study, was using the and ajusting the distribuition of YCBCR (for this case), and then generates the code automatically. The code created by the app can be read next.

```
function [BW,maskedRGBImage] = weedfilter1(RGB)
%createMask  Threshold RGB image using auto-generated code from colorThresholder app.
%  [BW,MASKEDRGBIMAGE] = createMask(RGB) thresholds image RGB using
%  auto-generated code from the colorThresholder App. The colorspace and
%  minimum/maximum values for each channel of the colorspace were set in the
%  App and result in a binary mask BW and a composite image maskedRGBImage,
%  which shows the original RGB image values under the mask BW.


% Auto-generated by colorThresholder app on 14-Feb-2017
%------------------------------------------------------------


% Convert RGB image to chosen color space
I = rgb2ycbcr(RGB);


% Define thresholds for channel 1 based on histogram settings
channel1Min = 67.000;
channel1Max = 87.000;


% Define thresholds for channel 2 based on histogram settings
channel2Min = 120.000;
channel2Max = 140.000;


% Define thresholds for channel 3 based on histogram settings
```

```
channel3Min = 102.000;
channel3Max = 120.000;


% Create mask based on chosen histogram thresholds
sliderBW = (I(:,:,1) >= channel1Min ) & (I(:,:,1) <= channel1Max) & ...
    (I(:,:,2) >= channel2Min ) & (I(:,:,2) <= channel2Max) & ...
    (I(:,:,3) >= channel3Min ) & (I(:,:,3) <= channel3Max);
BW = sliderBW;


% Initialize output masked image based on input image.
maskedRGBImage = RGB;


% Set background pixels where BW is false to zero.
maskedRGBImage(repmat(~BW,[1 1 3])) = 0;


end
```
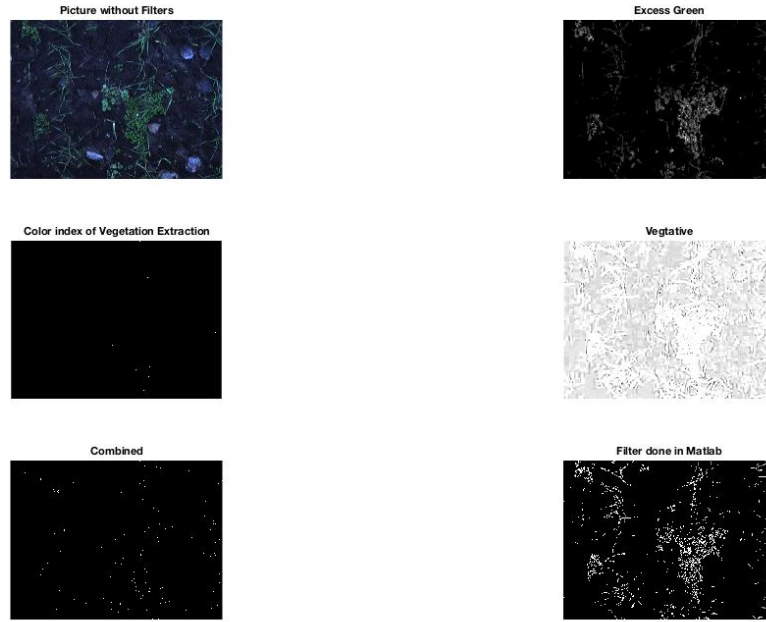
Figure A.1: All color indexes

A.2.2   Color Indexes

The color indices's used were the sames used in Support Vector Machines for crop/weeds identification in maize fields. In a RGB image, every pixel has three components, Red(R), Green(G) and Blue(B), that vary between 0-255. so first the three components are normalized:

$$R_n = \frac{R}{255} \qquad G_n = \frac{G}{255} \qquad R_n = \frac{B}{255} \tag{A.2}$$

$$r = \frac{R}{R_n + G_n + B_n} \qquad g = \frac{G}{R_n + G_n + B_n} \qquad b = \frac{B}{R_n + G_n + B_n} \tag{A.3}$$

The excess green(ExG):

$$ExG = 2g - r - b \tag{A.4}$$

The color index of vegetation extraction($CIVE$):

$$CIVE = 2g - r - b \tag{A.5}$$

The VEG(VEG):

$$VEG = b = \frac{g}{r^a + b^{1-a}} \qquad with\,a\,set\,to\,0,667 \tag{A.6}$$

The combinations of the three(COM):

$$COM = 0.35ExG + 0.47CIVE + 0.17VEG; \tag{A.7}$$

A.2.3 Red line emitted by the NDVI sensor

The green seeker emitted a red light, which is the visual red light, VIS in eq.1.1. This light per-mitted to be sure that the camera was in the same direction of the NDVI. In fig.A.2
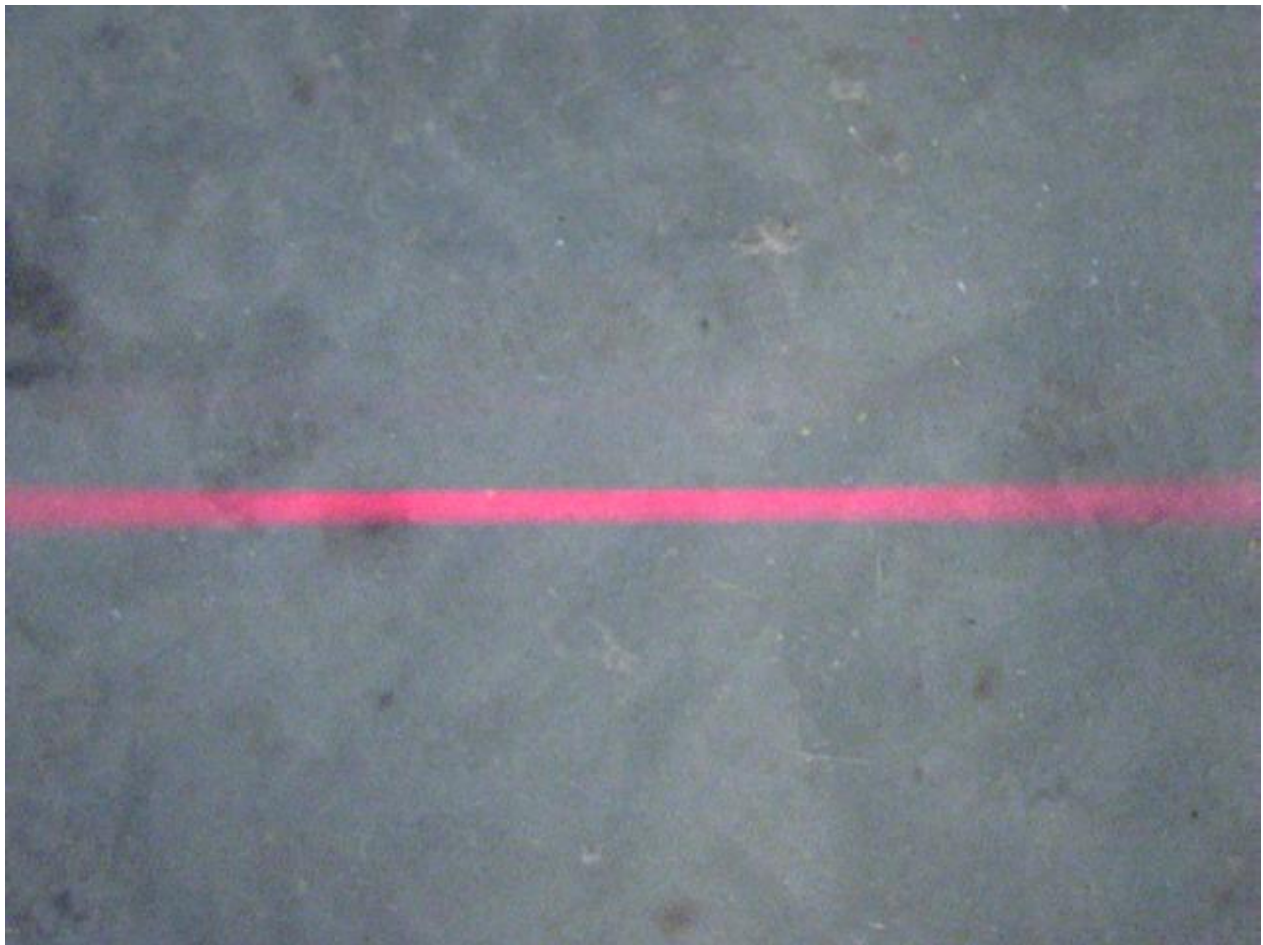

Figure A.2: Red beam emited by the GreenSeeker

Appendix B
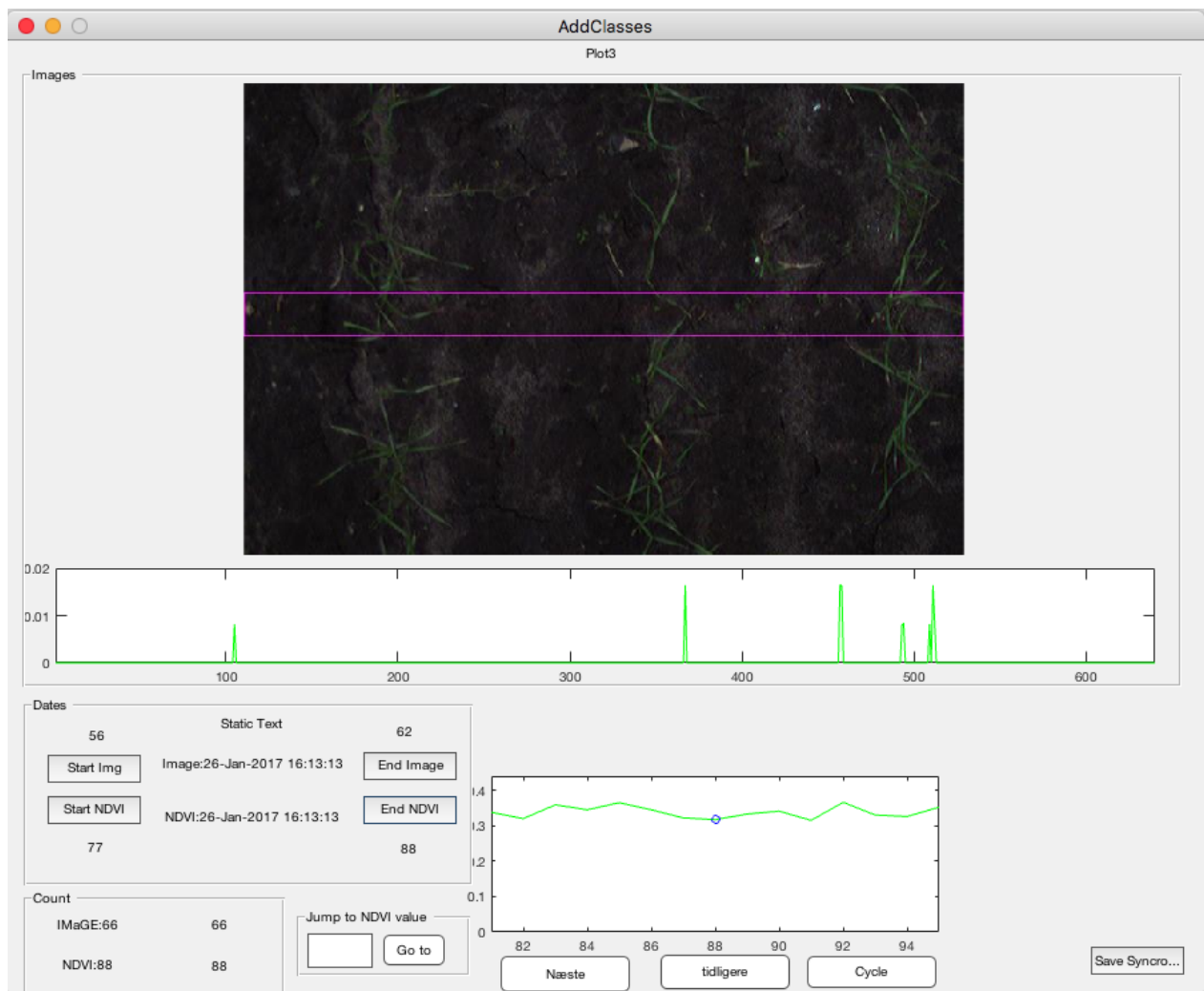
B.1 Syncronizations of Data

Figure B.1: Print screen of the application to synchronize the images with the NDVI

In fig.B.2 a print screen of the application to synchronize the images and the NDVI. The application was build specially for this study. in the when the NDVI started passed the white board(which would show in the plot in the bottom of the GUI) the user clicked the button "Start NDVI", and when the image was "leaving" the white board the used clicked the "Start Image", and the same for the End's buttons. The "cycle button" is for when a measurement is done, and the quadbike is turning or stopping. The plot bellow the image is the sum of the green values of each column of pixels.

B.2   Classification of Images

In fig.B.3 the print screen of the application used to classify the images.The application was build specially for this study. Depending on the image showing, the user pushes one of the "C" buttons, being C1 for low crop, C2 low crop with weed patch, C3 for High crop and C4 for high crop with weed patch. The "No weeds" or "weeds" was for when it was not clear if it was high crop or low crop.
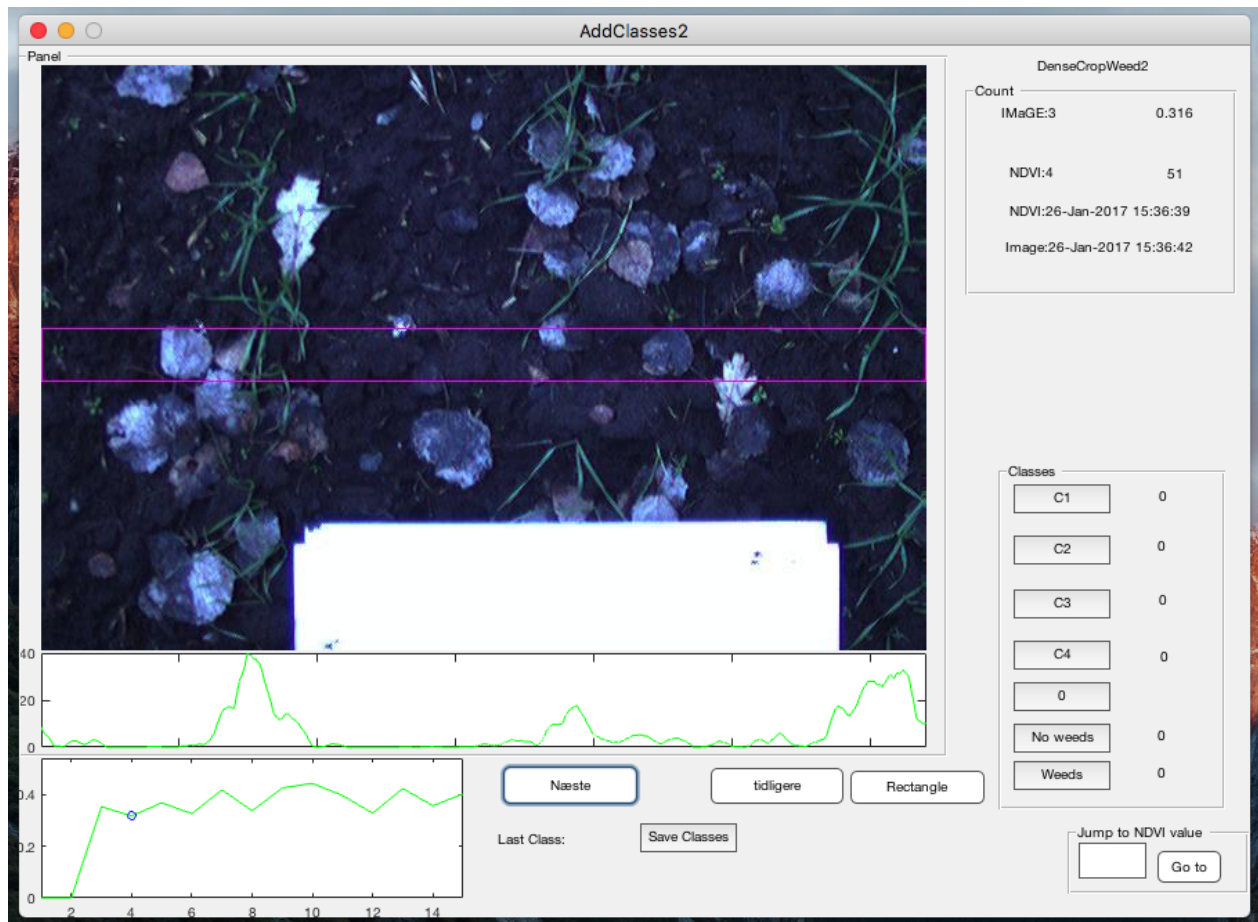

Figure B.2: Print screen of the application to synchronize the images with the NDVI

B.3   Example of problem in time stamp

There were parts were the camera didn't save the images in the regular interval, creating a big interval in the time stamp, causing the seconds to overlap when the fraction of seconds was added. Even thought, the classification keeps synchronized with time, this wouldn't happen in the log order.
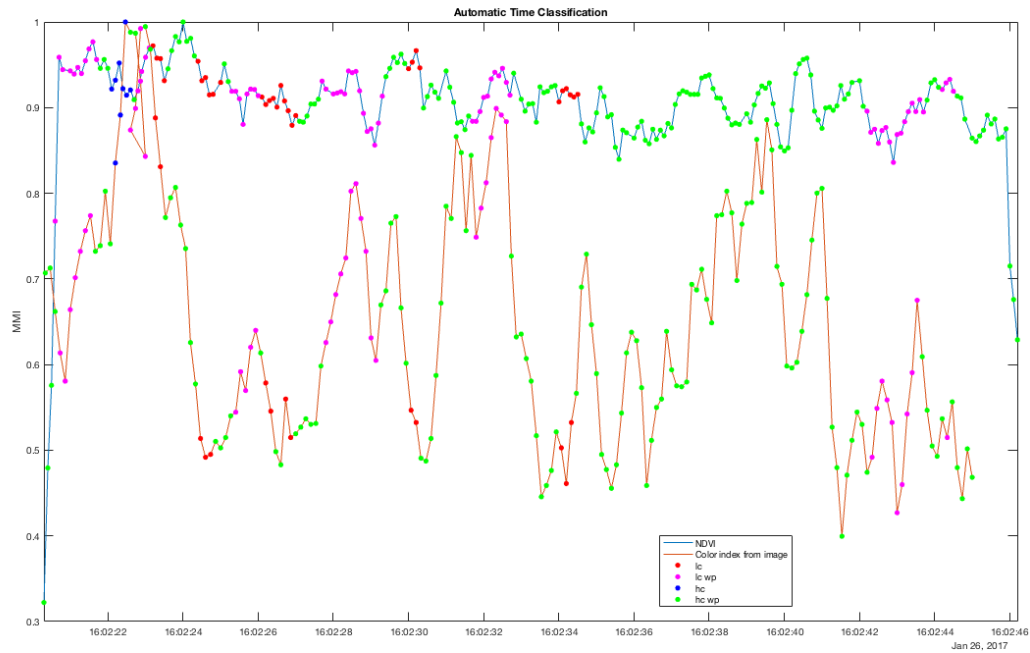

Figure B.3: Time Stamp Error in the seconds 22 to 24, on top is the NDVI by log order, in the middle, this are measurements in plot 5

Bibliography

(2009). Bekæmpelsesmiddelstatistik. (4).

Ahmed, F., Al-Mamun, H. A., Bari, A. S. M. H., Hossain, E., and Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection*, 40:98–104.

Ali, A. M., Thind, H. S., and Sharma, S. (2014). Field Crops Research Prediction of dry direct-seeded rice yields using chlorophyll meter, leaf color chart and GreenSeeker optical sensor in northwestern India. *Field Crops Research*, 161:11–15.

De Rainville, F. M., Durand, A., Fortin, F. A., Tanguy, K., Maldague, X., Panneton, B., and Simard, M. J. (2012). Bayesian classification and unsupervised learning for isolating weeds in row crops. *Pattern Analysis and Applications*, pages 1–14.

Foody, G. M. and Mathur, A. (2004). Toward intelligent training of supervised image classifications : directing training data acquisition for SVM classification. 93:107–117.

Guerrero, J. M., Pajares, G., Montalvo, M., Romeo, J., and Guijarro, M. (2012). Expert Systems with Applications Support Vector Machines for crop / weeds identification in maize fields. *Expert Systems With Applications*, 39(12):11149–11155.

Hong, S., Minzan, L., and Zhang, Q. (2012). Detection system of smart sprayers : Status , challenges , and perspectives. 5(3):10–23.

MIT OpenCourseWare (2014). Learning: Support Vector Machines.

Press, A. (2017). Selecting Appropriate Weed Control Systems for Developing Countries Author (s): Douglas Young , Stanley Miller , Herbert Fisher and Myron Shenk Published by : Weed

Science Society of America and Allen Press Stable URL : http://www.jstor.org/stable/40427. 26(3):209–212.

Rajab, S., Kalaa, M. O. A., and Refai, H. (2016). Classi fi cation and speed estimation of vehicles via tire detection using single-element piezoelectric sensor. (September):1366–1385.

Ruiz-gonzalez, R., Gomez-gil, J., Gomez-gil, F. J., and Martínez-martínez, V. (2014). An SVM-Based Classifier for Estimating the State of Various Rotating Components in Agro-Industrial Machinery with a Vibration Signal Acquired from a Single Point on the Machine Chassis. (i):20713–20735.

Salappa, A., Doumpos, M., and Zopounidis, C. (2017). Feature selection algorithms in classification problems : an experimental evaluation. 6788(March).

Samborski, S. M., Gozdowski, D., Walsh, O. S., Lamb, D. W., Stępień, M., Gacek, E. S., and Drzazga, T. (2015). Winter Wheat Genotype Effect on Canopy Refl ectance: Implications for Using NDVI for In-Season Nitrogen Topdressing Recommendations. (May):2097–2106.

San, C., Andújar, D., Barroso, J., Fernández-quintanilla, C., and Mart, C. S. Weed Decision Threshold as a Key Factor for Herbicide Reductions in Site- Specific Weed Management Weed Decision Threshold as a Key Factor for Herbicide Reductions in Site- Specific Weed Management. 30(4):888–897.

Swain, K. C., Nørremark, M., Jørgensen, R. N., Midtiby, H. S., and Green, O. (2011). Weed identification using an automated active shape matching (AASM) technique. *Biosystems Engineering*, 110(4):450–457.

Tellaeche, A., Pajares, G., Burgos-artizzu, X. P., and Ribeiro, A. A computer vision approach for weeds identification through Support Vector Machines.

Tellaeche, A., Pajares, G., Burgos-artizzu, X. P., and Ribeiro, A. (2011). A computer vision approach for weeds identification through Support Vector Machines. 11:908–915.

Terzic, J., Nagarajah, C. R., and Alamgir, M. (2010). Sensors and Actuators A : Physical Fluid level measurement in dynamic environments using a single ultrasonic sensor and Support Vector Machine (SVM). 161:278–287.

Weier, J., Herring, D., Difference, N., and Index, V. (2006). NDVI and EVI. pages 1–7.

Zheng, B., Myint, S. W., Thenkabail, P. S., and Aggarwal, R. M. (2015). International Journal of Applied Earth Observation and Geoinformation A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. *International Journal of Applied Earth Observations and Geoinformation*, 34:103–112.