

Norwegian University  
of Life Sciences

**Master's Thesis 2023 30 ECTS**

Faculty of Science and Technology (REALTEK)

# **Detection of Fusarium Head Blight in Wheat Grains Using Hyperspectral and RGB Imaging**

**Daniel Stebekk Leirdal**

Data Science

## **Acknowledgement**

First and foremost, I would like to express my sincere gratitude to my thesis supervisor Prof. Ingunn Burud and co-supervisor Sahameh Shafiee for their invaluable support and guidance throughout the research process. Their willingness to answer my questions and provide me with the necessary resources and literature has been instrumental in the completion of this study. I am grateful for their dedication and expertise, which have helped to keep me on track and focused on my research goals. Further, I would also like to thank Prof. Morten Lillemo for the guidance involving the data acquisition and preparation of the samples.

I would also like to express my appreciation to my family for their consistent support throughout this thesis journey. Their encouragement, love, and belief in my abilities have been a constant source of motivation. Their patience and understanding during the times when I needed to work long hours and the times when I felt stressed have been invaluable. I am truly grateful for their sacrifices and for always being there to offer a listening ear.



## ABSTRACT

In modern agriculture, it is imperative to ensure that crops are healthy and safe for consumption. Fusarium Head Blight (FHB) can cause significant damage to wheat grains by reducing essential components such as moisture, protein, and starch, while also introducing dangerous toxins. Therefore, accurately distinguishing between healthy and FHB-infected wheat grains is essential to guarantee stable and reliable wheat production while limiting financial losses and ensuring food safety. This thesis proposes effective methods to classify healthy and FHB infected wheat grains using Hyperspectral Imaging (HSI) and Red Green Blue (RGB) images. The approach includes a combination of Principal Component Analysis (PCA) with morphology, in addition to dark and white reference correction, to create masks for grains in each image. The classification for the hyperspectral images was achieved using a Partial Least Squares Discriminant Analysis (PLS-DA) model for hyperspectral images and a Convolutional Neural Network (CNN) model for RGB images. Both object-based and pixel-based approaches were compared for the PLS-DA model. The results indicated that the object-based approach outperformed the pixel-based approach and other well-known machine learning algorithms, including Random Forest (RF), linear Support Vector Machine (SVM), Stochastic Gradient Descent (SGD) calibrated one-vs-all and DecisionTree. The PLS-DA model using the object-based method yielded better results when tested on all wheat varieties, achieving an F1-score of 99.4%. Specific wavelengths were investigated based on a loading plot, and four effective wavelengths were identified, 953 *nm*, 1373 *nm*, 1923 *nm* and 2493 *nm*, with classification accuracy found to be similar to the full spectral range. Moreover, the moisture and water content in the grains were analyzed using hyperspectral images through an aquagram, which demonstrated that healthy grains exhibited higher absorbance values than infected grains for all Water Matrix Coordinates (WAMACS). Furthermore, the CNN model was trained on cropped individual grains, and the classification accuracy was similar to the PLS-DA model, with an F1-score of 98.1%. These findings suggest that HSI is suitable for identifying FHB-infected wheat grains, while RGB images may provide a cost-effective alternative to hyperspectral images for this specific classification task. Further research should consider to explore the potential benefits of HSI for deeper investigations into how water absorption affects spectral measurements and moisture content in grains, in addition to user-friendly interfaces for deep learning based image classification.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Power of Wheat - Driving Agricultural Innovation . . . . .	1
1.2	Impacts of FHB in Agriculture . . . . .	1
1.3	Machine Learning-Based Detection of Fungal Infections . . . . .	2
1.4	Related Works . . . . .	2
1.5	Outline of Thesis Structure . . . . .	3
<b>2</b>	<b>Theory</b>	<b>5</b>
2.1	Visual Data Analysis . . . . .	5
2.2	Hyperspectral Sensing . . . . .	6
2.2.1	Electromagnetic Spectrum and Hypercube . . . . .	6
2.2.2	Properties of Molecules and Their Interaction with Light . . . . .	7
2.2.3	Imaging Spectroscopy - Principle of Operation . . . . .	8
2.3	Machine Learning Algorithms . . . . .	9
2.3.1	Chemometrics . . . . .	10
2.3.2	Partial Least Squares . . . . .	10
2.3.3	Artificial Neural Networks . . . . .	12
2.3.4	Convolutional Neural Networks . . . . .	13
2.3.5	Activation and Loss Functions . . . . .	14
2.3.6	Metrics for Binary Classification . . . . .	16
2.4	Breeze Software . . . . .	17
2.5	Aquaphotomics . . . . .	18
<b>3</b>	<b>Methodology</b>	<b>20</b>
3.1	Data Collection . . . . .	20
3.1.1	Grains . . . . .	20

3.1.2	Image Acquisition . . . . .	21
3.2	Calibration and Preprocessing Techniques . . . . .	23
3.3	Breeze Workflow for Spectral Imaging . . . . .	27
3.3.1	PCA - Feature Extraction and Segmentation . . . . .	27
3.3.2	PLS-DA - Classification . . . . .	28
3.3.3	Effective Wavelengths . . . . .	29
3.4	Predicting Grain Condition using CNN with RGB Images . . . . .	29
3.4.1	Mask Generation and Preparation of Dataset . . . . .	29
3.4.2	Parameter Optimization . . . . .	31
3.4.3	Data Augmentation . . . . .	31
3.4.4	Data Distribution . . . . .	31
3.4.5	CNN Architecture . . . . .	32
3.5	Exploring Alternative Machine Learning Approaches . . . . .	34
3.6	Aquaphotomics Analysis of Wheat Grains . . . . .	34
<b>4</b>	<b>Results</b>	<b>36</b>
4.1	Classification of the Hyperspectral Data . . . . .	36
4.1.1	Score Variance Plot Analysis . . . . .	36
4.1.2	Object-Based vs. Pixel-Based Classification . . . . .	37
4.1.3	All Wavelengths vs. Effective Wavelengths . . . . .	38
4.2	Deep Learning and Non-Deep Learning Classification . . . . .	41
4.2.1	CNN-based Analysis with RGB Images . . . . .	41
4.2.2	Alternative Machine Learning Algorithms . . . . .	42
4.3	Moisture and Water Content with Aquaphotomics Analysis . . . . .	43
<b>5</b>	<b>Discussion</b>	<b>46</b>
5.1	Image Acquisition and Preprocessing . . . . .	46
5.2	Classification of Spectral Data . . . . .	46

5.2.1	Loadings . . . . .	47
5.2.2	Effect of Wheat Varieties on Grain Classification . . . . .	47
5.3	Classification of RGB Images . . . . .	48
5.4	Moisture and Water Content . . . . .	48
5.5	Limitations and Uncertainties . . . . .	49
<b>6</b>	<b>Conclusion</b>	<b>51</b>
<b>A</b>	<b>Appendix 1 - Limited Wheat Varieties in Test Set</b>	<b>60</b>
<b>B</b>	<b>Appendix 2 - Every Variety - Effective Wavelengths</b>	<b>76</b>
<b>C</b>	<b>Appendix 3 - Every Variety - All Bands</b>	<b>89</b>

# List of Figures

2.1	Classification, semantic segmentation, object detection, instance segmentation. Source: [1]. . . . .	5
2.2	The electromagnetic spectrum from ultraviolet to near-infrared on top, including a monochromatic image, stacked red, green, and blue channels and a final image on the right showcasing a hypercube of the presenting image. . . . .	6
2.3	Illustration of the different interactions between photons and matter. When incident light ( $I_0$ ) interacts with a sample, its trajectory and energy can be altered due to absorption ( $I_a$ ), reflection ( $I_R$ ), or transmission ( $I_T$ ) by the sample's molecules. . .	7
2.4	Principle of HySpex camera. From electromagnetic waves to hypercube. Figure inspired by HySpex. [Hyspex.com] . . . . .	8
2.5	General workflow of machine learning algorithms. Training data is fed into the model to make a prediction on new data. Figure inspired by [2]. . . . .	9
2.6	Scatter plot showing the relationship between the variables extracted through PCA and PLS-DA using the Iris dataset [link to dataset], demonstrating the reduction of dimensionality and correlation among variables. . . . .	11
2.7	Demonstration of a simple, fully connected artificial neural network. . . . .	12
2.8	Architecture of the CNN used for binary classification of wheat grains. The figure shows the feature extraction process through convolutional layers, followed by classification using fully connected layers with a sigmoid output. The figure is inspired by Developers Breach [link to website]. . . . .	13
2.9	Comparison of the Sigmoid and ReLU activation functions commonly used in neural networks. The Sigmoid function maps inputs to a smooth output range between 0 and 1, while the ReLU function maps inputs linearly to a range between 0 and positive infinity. . . . .	15
2.10	Example of a confusion matrix for binary classifiers. . . . .	16
2.11	General workflow of Breeze software from Prediktera AB [3]. . . . .	17
3.1	Preparation and annotation of samples in the NMBU workshop. . . . .	20
3.2	Setup for image acquisition in laboratory. . . . .	22
3.3	Mean spectra with shaded areas representing the standard deviation for reflectance and absorbance of all healthy and infected grains in the dataset. . . . .	24

3.4	Hyperspectral image of healthy grains. Showing a 2-dimensional image with the corresponding (SNV corrected) mean absorbance spectra for each of the 20 grains in the hyperspectral image. . . . .	25
3.5	Analysis tree of the workflow in Breeze. From segmentation to classification. The first PLS-DA model contains four PC's based on an object-based evaluation, whereas the other PLS-DA model contains four PC's based on a pixel-based classification. . . . .	27
3.6	Process of pixel segmentation based on a scatter plot. After selecting a wavelength interval and applying pretreatment, the top image displays six raw images including background pixels, while the middle image shows the ROI. The last image shows the six resulting images after background removal, along with their correlated scatter plot. . . . .	28
3.7	Loading plot from the PLS-DA model showing the corresponding wavelengths and their importance for binary classification of the grains. The four specific wavelengths highlighted in red indicates a strong negative correlation between the two classes. . . . .	29
3.8	Process of generating binary mask using a combination of LAB color space and PCA, followed by thresholding and morphology operations. . . . .	30
3.9	Mask for RGB images with one label for each grain color-coded and their corresponding bounding boxes highlighted in red. . . . .	30
3.10	Distribution of the training, validation- and test set. . . . .	32
3.11	A schematic representation of the CNN architecture used for the binary classification task. The model consists of two convolutional layers with max pooling and batch normalization, followed by a fully connected layer with dropout and a sigmoid output layer. . . . .	33
3.12	EMSC absorbance spectra applied before normalizing the values to be plotted in aquagram. . . . .	35
4.1	Score variance plot to visualize the distribution of data points in the PLS-DA model for healthy and infected grains, and to identify potential outliers. . . . .	36
4.2	Confusion matrix for object-based classification using AMS on test data. . . . .	37
4.3	Confusion matrix for pixel-based classification using PM on test data. . . . .	37
4.4	Classification examples of infected grains using object-based and pixel-based methods. The PM classifier made incorrect classifications, while the AMS classifier correctly classified the chosen grains. . . . .	38

4.5	The presented images are composed of various spectral bands, wherein each image corresponds to a specific wavelength within four distinct intervals in addition to the RGB channels. The images on the left represent healthy grains, while those on the right show infected grains. . . . .	39
4.6	Confusion matrix for object-based classification using AMS with all bands on every wheat variety as test data. . . . .	39
4.7	Confusion matrix for object-based classification using AMS with EWs on every wheat variety as test data. . . . .	40
4.8	Examples of grain classification using EWs and all bands for every wheat variety. The "Healthy or Infected" column represents the ground truth, while the last column displays the prediction. . . . .	40
4.9	Accuracy and loss curves of the CNN model for training and validation datasets. . .	41
4.10	Confusion matrix predicted by the CNN model for binary classification of wheat grains. . . . .	42
4.11	Comparison of DecisionTree (LGBM), RF, one-vs-all SGD, and linear SVM models. Confusion matrices depict the test set results with corresponding precision, recall, and F1-score. . . . .	43
4.12	Second derivative spectra for healthy and infected grains with their corresponding WAMACS in each region. The dashed vertical line shows the chosen wavelength in each region. . . . .	44
4.13	Aquagram with WASPs for each WAMAC and pure water. . . . .	44
A.1	Object-based vs. Pixel-based classification with test set containing only a few wheat varieties. Page 1/16 . . . . .	60
B.1	Object-based classification using EWs with test set containing every wheat variety. The third column "Healthy or Infected" is the true label, while the predicted label is under the fourth column "EWs-PLSDA_varieties". Page 1/13 . . . . .	76
C.1	Object-based classification using all bands with test set containing every wheat variety. The third column "Healthy or Infected" is the true label, while the predicted label is under the fourth column "PLSDA_varieties". Page 1/13 . . . . .	89

## List of Tables

2.1	The table presents the ranges of WAMACS in the first overtone, their corresponding wavelengths, descriptions and assignment & vibration types. Including $S_{0-4}$ , hydration shell, stretching and bending. . . . .	19
3.1	Overview of the 20 different wheat varieties and their concentration of Fusarium Damaged Kernels (FDK) in percentage and deoxynivalenol (DON) contamination.	21
3.2	Table representing the hyperparameters used for optimization when building the CNN, with a total of 96 possible combinations. . . . .	31
3.3	Shape of the train-, validation- and test data . . . . .	32
4.1	Precision, recall, and F1 score for the CNN model based on the confusion matrix Fig. 4.10. . . . .	42



# List of Abbreviations and Acronyms

**AMS** Average Mean Spectrum

**ANNs** Artificial Neural Networks

**CNN** Convolutional Neural Network

**DA** Discriminant Analysis

**DNN** Deep Neural Network

**EMSC** Extended Multiplicative Signal Correction

**FHB** Fusarium Head Blight

**FN** False Negative

**FOV** Field of View

**FP** False Positive

**HSI** Hyperspectral Imaging

**ML** Machine Learning

**NIR** Near Infrared

**OLS** Ordinary Least Squares

**PCA** Principal Component Analysis

**PLS** Partial Least Squares

**PM** Pixel Majority

**RGB** Red Green Blue

**SNV** Standard Normal Variate

**TN** True Negative

**TP** True Positive

# 1 Introduction

## 1.1 The Power of Wheat - Driving Agricultural Innovation

As of today, wheat (*Triticum aestivum L.*) holds a prominent position in global agriculture. Human nutrition is closely related to grain consumption; wheat is one of the basic pillars of the human diet along with maize and rice, and about 35% of human's calories intake comes from these crops [4]. Production of these three grains has increased greatly over the last 6 decades [5]. In the ensuing decades, there will likely be a significant rise in the demand for wheat due to population growth, shifting dietary trends, and the growing use of wheat in the production of biofuels [6]. As a result, wheat farmers will face even more pressure to increase crop production while maintaining sustainability. Due to an increase in the frequency and severity of droughts, floods, and other extreme weather events, climate change is likely to exacerbate the problems with wheat production. Researchers are working to develop new wheat varieties that are more resistant to diseases and pests, as well as more tolerant of drought and other environmental stresses [7, 8]. This could help to improve yields and reduce losses due to factors like Fusarium Head Blight (FHB).

## 1.2 Impacts of FHB in Agriculture

FHB is a devastating fungal disease that poses a significant threat to wheat agriculture worldwide. It can cause substantial reductions in both the quantity and quality of grain, leading to yield losses of up to 40% [9]. The invasion of wheat kernels by Fusarium fungus and subsequent decomposition of storage proteins and starch within the pericarp leads to significant losses of water, protein, and starch content, resulting in the withering and shrinkage of the grains. This ultimately leads to high crop failure and a reduction in quality [10]. In addition, FHB can produce harmful mycotoxins such as deoxynivalenol (DON), which can contaminate food and make it inedible for human or animal consumption [11]. The conventional methods of detecting FHB predominantly rely on visual interpretation by trained personnel or employment of various chemical techniques such as gas chromatography (GC), high-performance liquid chromatography (HPLC), enzyme-linked immunosorbent assay (ELISA), and polymerase chain reaction (PCR) to identify the presence of FHB and the production of the mycotoxin DON. However, these methods are characterized by their time and labor-intensive nature, their inability to facilitate large-scale monitoring and their destructive effects on wheat [12, 13, 14, 15]. FHB causes economic losses of billions of dollars each year and can have a significant impact on food security, especially in developing countries [16]. Factors such as weather conditions, cropping systems, and fungicide use can influence the prevalence and severity of FHB outbreaks. Therefore, as previously stated, developing FHB-resistant wheat varieties is therefore crucial for ensuring agricultural productivity and food safety.

### 1.3 Machine Learning-Based Detection of Fungal Infections

To minimize crop loss due to fungal infections such as FHB, the use of Machine Learning (ML) algorithms for easier detection of infections has been proposed. One promising approach is Hyperspectral Imaging (HSI), which enables the study of wavelengths outside the visible range, specifically in the Near Infrared (NIR) range. By extracting spectral information from each pixel in every band, it is possible to spot specific spectral characteristics in the grains that indicate the presence of infections. Hyperspectral images, which contain a large number of bands, are therefore useful for capturing more spectral information. However, noise can affect the results when analyzing the pixel values and differentiating the grains. To address this issue, preprocessing techniques for hyperspectral images are commonly used. By applying ML algorithms to hyperspectral images, the detection of fungal infections in grains can be automated and made more accurate. Deep learning computer vision techniques, particularly Convolutional Neural Network (CNN), have proven to be effective for analyzing grain images under varying field conditions and diverse plant phenotypes. This approach is considered state-of-the-art for image classification and segmentation [17]. However, the primary challenge in training a CNN to accurately distinguish between healthy and infected grains under different conditions lies in the extensive amount of annotated training data required. Therefore, the data must span a wide range of natural variability, including images of both healthy and infected kernels [18]. The manual annotation of thousands of images is a time-consuming and costly process.

### 1.4 Related Works

The development of ML has already been extended to FHB detection and there exists a significant amount of research on the topic, particularly with regards to wheat grains. Qiu et al. [19] conducted a study utilizing a Deep Neural Network (DNN) with transfer learning, in conjunction with color imaging, for the detection of FHB-affected wheat spikes. The method for detection and segmentation of wheat spikes employed a region growing technique, whereby a pixel is initially selected as the seed point, and its surrounding pixels are compared and evaluated with it. Additionally, image augmentation was applied to increase the amount of data available for training the DNN, with a total of 2829 images utilized. Results indicated that the DNN, in addition to a Green Blue color feature to highlight the diseased areas of each spike, achieved a mean average precision score of 92% and  $R^2 = 80\%$ . It is noteworthy that limitations in the study included blurred images and a limited Field of View (FOV) to detect spikes.

Another study from Yipeng et al. [10] also represented a DNN, but with the use of monochromatic images. The model was trained to distinguish between sound, mildly, moderately and severely damaged wheat kernels. Effective wavelengths were chosen and the model received an accuracy of 100% and 98.31% in the training and prediction sets, respectively.

Zhao et al. [20] used metabolomics to analyze the differences in metabolic profiles between two wheat cultivars with different levels of resistance to FHB caused by *Fusarium graminearum*. The researchers found that specific amino acids, such as proline and alanine, increased resistance to FHB,

while cysteine aggravated susceptibility. Partial Least Squares - Discriminant Analysis (PLS-DA) was used to validate the data, with VIP values of greater than 1 identifying significant metabolite differences. This research provides valuable insights into the metabolite differences of two wheat cultivars in response to *F. graminearum* infection and highlights the potential of using metabolomics and PLS-DA for exploring the mechanisms involved in FHB resistance.

In addition to the use of metabolomics and PLS-DA in exploring the mechanisms involved in FHB resistance, there have been interesting applications of aquaphotomics in food science. For instance, a study from Malegori et al. [21] investigated the relationship between water activity and the molecular structure of rice germ. The study found that the state of the rice germ is influenced by the water activity during storage, which is defined by the structure of water within the germ matrix. The absorbance spectral pattern at several absorbance bands was found to be descriptive of the water structure.

Another study from Atanassova et al. [22] aimed to investigate the feasibility of using NIR and aquaphotomics to monitor changes during the ripening of Bulgarian yellow cheese from cow milk. The study analyzed samples at various stages of ripening and found significant changes in the aquagram patterns, which were related to changes in titratable acidity and protein fraction. The study suggests that NIR spectroscopy and aquaphotomics have potential as tools for monitoring cheese ripening stages.

Although aquaphotomics has been extensively studied in fields like food science and medicine, its application in the analysis of wheat grains has been relatively underexplored. This research gap presents an opportunity to explore the potential for further research on the use of aquaphotomics as a tool for detecting FHB infection and monitoring the quality of wheat grains.

## **1.5 Outline of Thesis Structure**

This thesis aims to develop a method for automatic detection of healthy or FHB infected grains using hyperspectral images with spectral information in the NIR range, in comparison to low-cost and more computationally efficient RGB images. To this end, this thesis proposes unsupervised and supervised models to classify objects and their corresponding pixel values effectively. Furthermore, this study explores the performance of classifiers and deep learning models for binary classification of wheat grains based on RGB and hyperspectral images using Pixel Majority (PM) and Average Mean Spectrum (AMS) classification methods.

Moreover, this study considers the relatively new field of aquaphotomics, which focuses on the use of infrared spectroscopy to investigate the interactions between water and biological samples. Wheat grains contain a considerable amount of water, and aquaphotomics can provide valuable information about their quality, including moisture content. Hence, aquaphotomics offers a promising tool for detecting FHB infection in wheat grains, as the spectral characteristics of water can reflect changes in the grain's chemical composition and structure caused by the disease.

To explore the potential of aquaphotomics for FHB detection, this study creates an aquagram from

Extended Multiplicative Signal Correction (EMSC) processed spectra with Standard Normal Variate (SNV) correction to display certain wavelengths that could be affected by water content in the grains. By incorporating aquaphotomics as a tool, this study provides a novel approach to FHB detection in wheat grains.

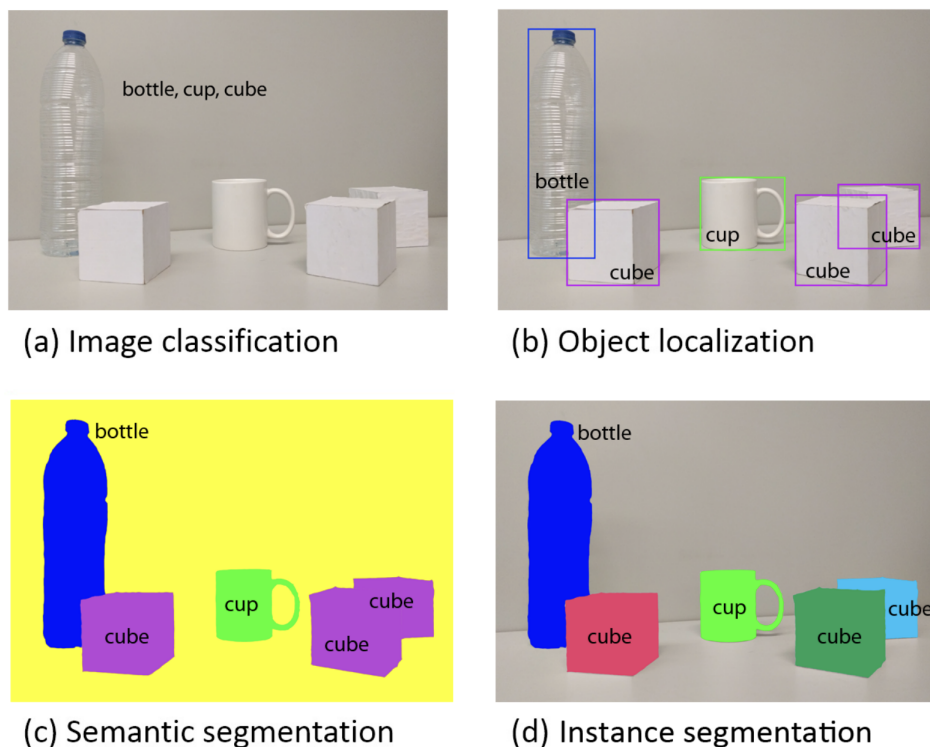
The purpose of this study was to investigate the following research questions:

- How do classifiers and deep learning models perform in the classification of healthy or FHB-infected wheat grains based on RGB and hyperspectral images using pixels-based and object-based methods?
- Can RGB images replace hyperspectral images for binary classification of wheat grains, and what is their accuracy compared to hyperspectral images?
- How is the moisture and water content of healthy and infected grains, represented through aquagram?

## 2 Theory

### 2.1 Visual Data Analysis

Visual data analysis is a field that aims to allow machines to have the ability to analyze and interpret image data in a similar manner to human beings [23]. A data analysis system comprises images and videos to extract information along with additional data like timestamps, camera settings, or sensor readings [23]. The system's output may take different forms, such as an enhanced representation of the input image, or a "decision" based on the image content, such as identifying or counting objects, like seeds or grains. The primary goal of data analysis is to replace human vision in various challenging tasks, such as medical image analysis, autonomous driving, and quality control in industrial production. However, these tasks pose significant challenges for computers, which are better suited to processing static concepts and forms rather than abstract concepts that can take on multiple forms [18]. Visual data analysis can be classified into four main types of tasks: image classification, semantic segmentation, object detection, and instance segmentation [1]. These tasks involve identifying the content of an image, categorizing pixels into various classes, locating objects within an image, and determining the boundaries of individual objects, see Fig. 2.1. The key point of segmentation is to identify the most essential and valuable data by categorizing a larger set of data into smaller and more manageable segments based on specific criteria. While this process may involve excluding redundant data, the main focus is on grouping the data according to their relevant characteristics [24].



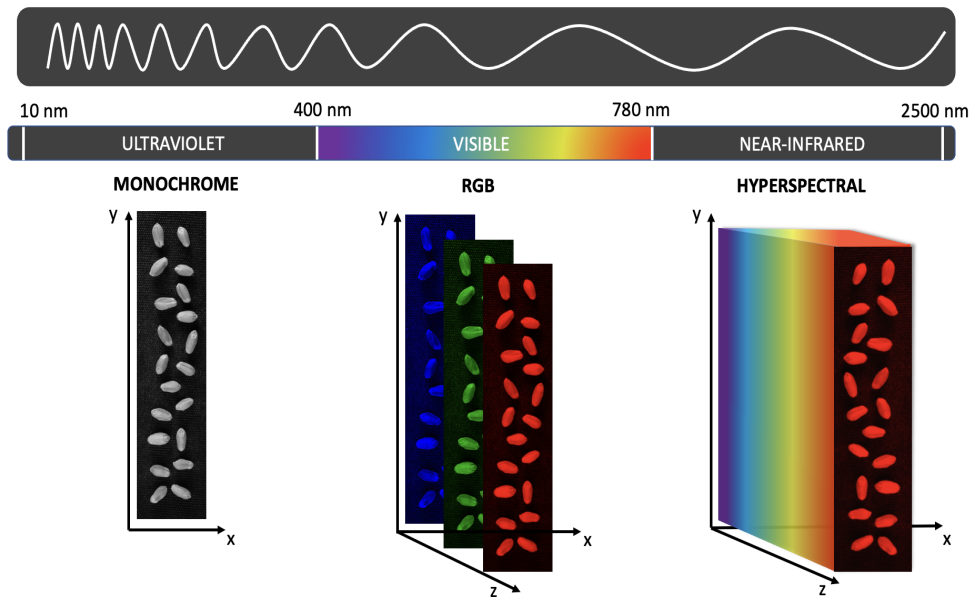
**Fig. 2.1:** Classification, semantic segmentation, object detection, instance segmentation.  
Source: [1].

## 2.2 Hyperspectral Sensing

Hyperspectral sensing is a powerful technique that involves using hyperspectral sensors to capture images with high spectral resolution. These images, referred to as hyperspectral images, are composed of hundreds of contiguous narrow wavelength bands, providing an abundance of spectral and spatial information about the imaged object [24]. Hyperspectral Imaging (HSI) has numerous applications in different fields, including agriculture, environmental monitoring, mineral exploration, and biomedical imaging, among others [25]. In agriculture, for example, HSI can be used to disease detection, monitor crop growth and yield prediction [26].

### 2.2.1 Electromagnetic Spectrum and Hypercube

The hyperspectral data can be represented as a hypercube, which is a three-dimensional array of data, with the first two dimensions (x- and y-axis) representing the spatial information, and the third dimension (z-axis) capturing the spectral information. Each pixel in the hypercube represents the continuous spectrum of light reflected or emitted by the corresponding location in the imaged object [27]. An example of a hypercube is shown in Fig. 2.2.

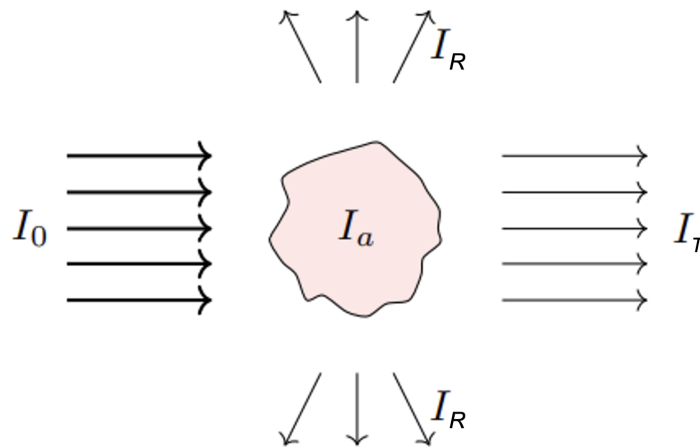


**Fig. 2.2:** The electromagnetic spectrum from ultraviolet to near-infrared on top, including a monochromatic image, stacked red, green, and blue channels and a final image on the right showcasing a hypercube of the presenting image.

Compared to conventional RGB images, hyperspectral images offer a much more extensive range of spectral information, capturing information from the ultraviolet to the infrared range [28]. This increased spectral resolution allows for the identification of specific features of the imaged object that are not visible to the naked eye. In contrast, RGB images only provide three primary colors, limiting the amount of information that can be captured [29].

## 2.2.2 Properties of Molecules and Their Interaction with Light

The interaction of photons with matter is a crucial concept in HSI. This interaction is determined by both the chemical properties of the molecules and the physical characteristics of the surface, such as roughness and compaction level. When a photon is emitted from a light source, it has a specific energy and trajectory that will be altered by its interaction with the sample. The energy of the photon will decrease due to absorption by the sample's molecules, and its trajectory will be altered [30]. This change in trajectory can result in the photon being completely absorbed, reflected, or transmitted through the sample (see Fig. 2.3).



**Fig. 2.3:** Illustration of the different interactions between photons and matter. When incident light ( $I_0$ ) interacts with a sample, its trajectory and energy can be altered due to absorption ( $I_a$ ), reflection ( $I_R$ ), or transmission ( $I_T$ ) by the sample's molecules.

Absorption occurs when the photon's frequency matches the vibrational frequency of the electrons in a molecule, causing the photon to be absorbed and the energy to be converted into heat [31]. One advantage of utilizing absorbance spectra is its resistance to the interference of light-scattering particles, such as dust or other contaminants, in contrast to reflectance spectra [32]. However, as grains are typically sensitive to the presence of water or other solvents, the use of reflectance spectra may be more appropriate for analyzing hydrated or wet samples that exhibit varying moisture content [32]. It is worth noting that the suitability of either spectral data depends on the specific application and sample type.

Reflection and transmission occur when the frequencies of the photon waves do not match the natural frequencies of the molecules in the sample. Reflection occurs when the photon is reflected off the surface of the sample, while transmission occurs when the photon is able to pass through the sample. If the sample is opaque to the frequencies of the photon, the photon will be reflected, allowing it to be detected.

Reflection can occur in a specular mode, where the photon is reflected at the same angle as the incident angle, or in a scattered mode, where the photon is reflected at a different angle [33]. In

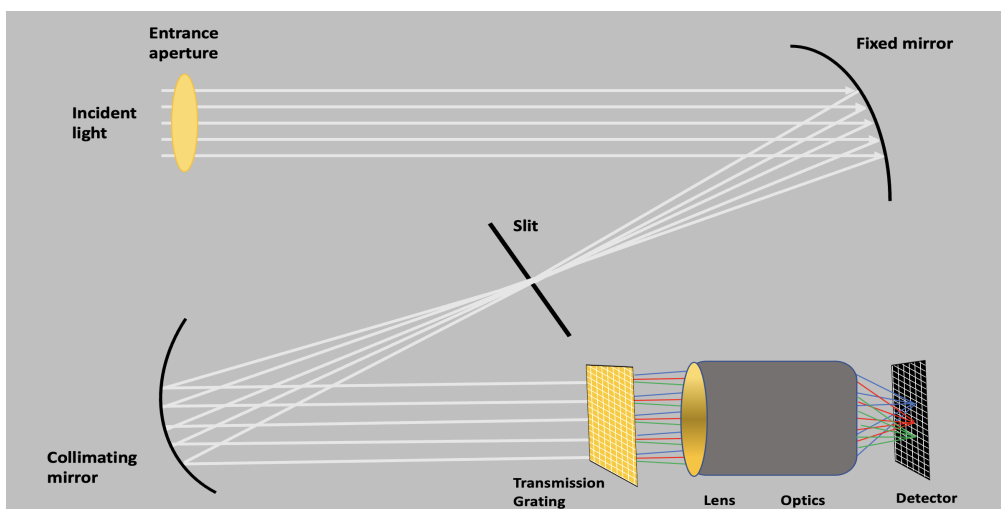


some cases, reflection may occur after the photon has penetrated the sample to a certain depth. The extent of penetration and the path the photon takes through the sample will depend on the photon's energy and the characteristics of the sample.

HSI techniques make use of the different interactions between light and matter to gather information about the composition and properties of a sample. This information can be used in a variety of fields, including remote sensing, medical diagnostics and agriculture. When analyzing grains as samples, the photons can penetrate a certain depth into the grain, depending on the energy of the photon and the grain's properties [33]. This information can provide insight into the chemical and physical properties of the grain.

### 2.2.3 Imaging Spectroscopy - Principle of Operation

Since hyperspectral images can obtain the individual color signature for each distinct object, HSI enables the ability to distinguish the full color spectrum in each pixel [27]. As a result, it consists of spectral information in addition to regular 2D RGB images. A feasible approach for acquiring a 2D image for each spectral channel when dealing with hyperspectral images involves scanning the scene in a sequential, line-by-line manner, with the aid of a pushbroom scanner, also referred to as a line-by-line scanner [34]. This type of hyperspectral camera captures incident light through a lens and uses a dispersive element, either a prism or diffraction grating, to split the light into different wavelengths. The camera operates by imaging the scene onto a slit that only allows light from a narrow line in the scene to pass through. After collimation, the transmission grating separates the different spectral lengths, and the light is focused onto a detector array [Hypspec.com]. The diffracted light will be captured by the detector array to read the information. The process is shown in Fig. 2.4. Each pixel interval along the slit corresponds to a spectrum that is projected on a column of detectors on the array, resulting in a hypercube with two spatial dimensions and one spectral dimension.



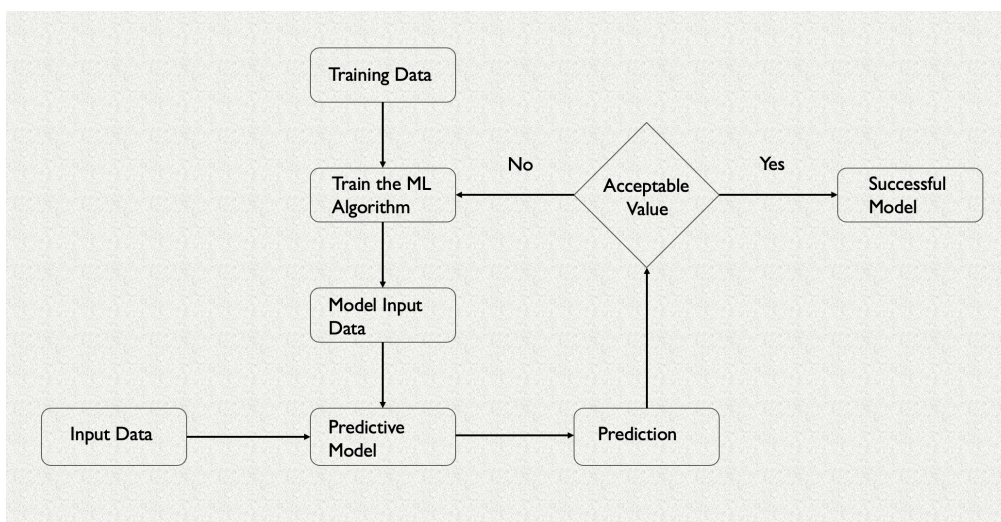
**Fig. 2.4:** Principle of HySpex camera. From electromagnetic waves to hypercube. Figure inspired by HySpex. [Hypspec.com]

## 2.3 Machine Learning Algorithms

Machine Learning (ML) is a field of study that focuses on creating algorithms that can automatically learn from data and make predictions or decisions without being explicitly programmed [35]. These algorithms run on computers and are designed to process large amounts of data, making it possible to derive patterns and knowledge that would be difficult or impossible to discover through manual analysis [35]. The significance of ML has grown substantially in the modern era, as it finds wide-ranging applications in diverse fields like email filtering, voice recognition, computer vision, and several others [36].

ML is a vast field that involves a diverse range of algorithms and techniques. Among the different types of ML algorithms, supervised, unsupervised, and reinforcement learning [35] are the most commonly used. Supervised learning involves training a model using labeled data, where the desired output signals (labels) are already known. This allows the algorithm to learn to predict outcomes on new, unseen data. In contrast, unsupervised learning works with unlabeled data and tries to identify meaningful structure or patterns within the data. This can involve grouping the data into clusters or performing dimensionality reduction for data compression. Reinforcement learning focuses on developing models that improve their performance through interactions with the surroundings. In this type of learning, the model receives feedback in the form of rewards or penalties based on its actions, allowing it to learn a series of actions that will lead to the desired outcome [35].

To illustrate how a general ML operates, Fig. 2.5 demonstrates how training data, either labeled or unlabeled, is fed into a model to predict certain outcome of unseen data [2]. Each type of algorithm has its unique characteristics and applications, making them suitable for different types of problems.



**Fig. 2.5:** General workflow of machine learning algorithms. Training data is fed into the model to make a prediction on new data. Figure inspired by [2].

### 2.3.1 Chemometrics

The following sections will provide a detailed exposition of various ML algorithms used in this thesis, including Partial Least Squares - Discriminant Analysis (PLS-DA), Principal Component Analysis (PCA) and Convolutional Neural Network (CNN). The goal of this exposition is to enhance the reader's comprehension of the operational mechanics and their ability to address various problems.

### 2.3.2 Partial Least Squares

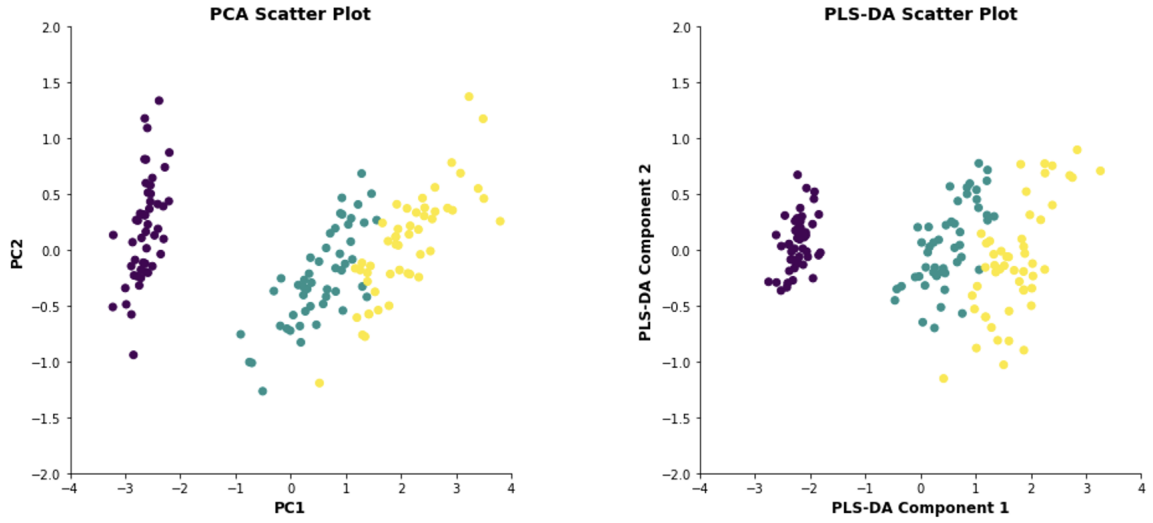
To provide some background information, it is worth noting that Partial Least Squares (PLS) is a technique that is closely related to Ordinary Least Squares (OLS), which is the traditional mathematical method used to fit a Linear Regression model [37].

Linear Regression is a statistical technique that aims to establish the relationship between a dependent variable and multiple independent variables [35]. To achieve this, OLS is commonly used, provided that the assumptions of Linear Regression are met. However, in some cases, the model may contain a large number of independent variables that are highly correlated with each other, resulting in multicollinearity and violation of the assumptions of Linear Regression [35].

On the other hand, PLS offers a potential solution to this problem by reducing the dimensionality of correlated variables and capturing the shared information among the variables, including the dependent and independent variables [37]. Additionally, PLS can model multiple outcome variables, which is a significant advantage compared to other statistical and ML models that are restricted to one outcome variable.

While alternative methods, such as building a model per variable, can be used for modelling multiple outcome variables, PLS is preferable in many analytical use cases because it enables the interpretation of a multivariate model, which differs from the interpretation of many univariate models. Therefore, PLS can be considered a powerful tool for dealing with complex datasets that contain multiple correlated variables and multiple outcome variables [38].

To address the issues of collinearity and dimensionality, multivariate methods like PCA and PLS-DA can be used. These methods work by extracting latent variables through linear combinations of the existing variables in the dataset, resulting in orthogonal variables that are less collinear [39]. These extracted variables can then be used in place of the original variables, significantly reducing the number of variables in the dataset. As shown in Fig. 2.6, the scatter plot illustrates the relationship between the variables extracted through PCA and PLS-DA.



**Fig. 2.6:** Scatter plot showing the relationship between the variables extracted through PCA and PLS-DA using the Iris dataset [link to dataset], demonstrating the reduction of dimensionality and correlation among variables.

PCA is an unsupervised method that extracts the latent variables based on the variation present in the data [40]. However, there is no guarantee that the selected components are relevant to the target variable, as the chosen components are obtained to explain the variance in the independent variables [40].

On the other hand, PLS-DA is a supervised method that works similarly to PCA but finds latent variables that are also relevant to the dependent variable [41]. PLS-DA performs the decomposition of both the independent and categorical dependent variables simultaneously, with the constraint that the selected components explain the maximum covariance between the two sets of variables [41]. After decomposition, a classification step is performed, where the decomposition of the independent variables is used to predict the class membership of the categorical dependent variable.

PLS-DA is particularly useful when dealing with datasets containing multiple correlated independent variables and a categorical dependent variable. By identifying the latent variables that are most relevant to the categorical dependent variable, PLS-DA allows for a more accurate prediction of the categorical dependent variable [42].

The underlying model can be explained by Eqs. (1 & 2).

$$X = TP^T + E_X \quad (1)$$

$$Y = UQ^T + E_Y \quad (2)$$

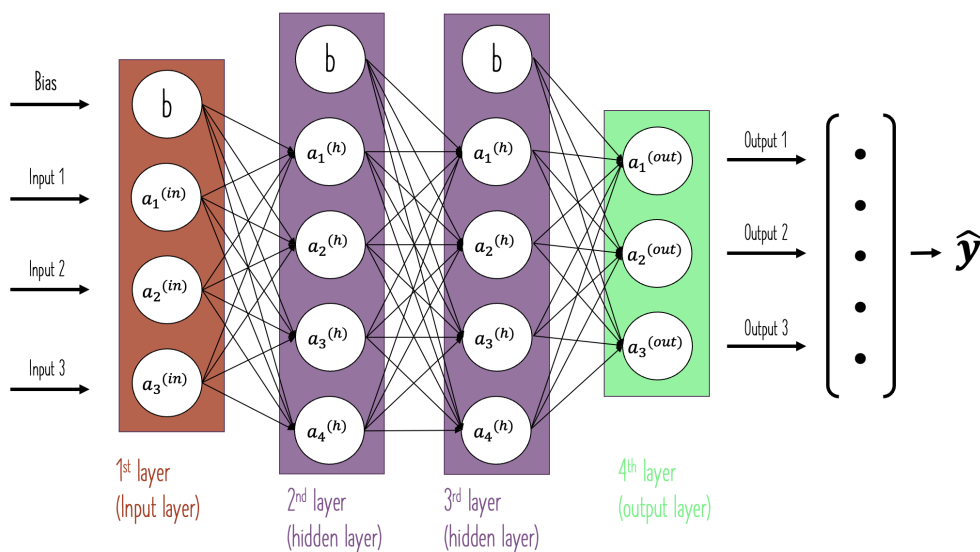
where  $T$  and  $U$  are scores matrices and  $P$  and  $Q$  are loading matrices of the target  $Y$  and data  $X$ , respectively. Additionally,  $E_X$  and  $E_Y$  correspond to the error terms [42].

### 2.3.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) have emerged as a powerful tool for solving complex classification and regression tasks, particularly in areas such as speech and image recognition [43]. ANNs are based on models of the human brain and are often referred to as "deep learning," reflecting the fact that they have a high number of layers and nodes, which enable them to learn and solve complex problems. ANNs are typically comprised of an input layer, one or more hidden layers, and an output layer, and are designed to learn from training data consisting of explanatory variables and response variables or labels [35].

While ANNs offer significant advantages over other algorithms in terms of their robustness and versatility, their performance is highly dependent on the quantity and quality of training data [44, 45]. Insufficient training data can lead to overfitting, where the network learns to model the patterns in the training data too well and performs poorly on new, unseen data [46]. As a result, it is crucial to have sufficient high-quality training data when using ANNs for complex tasks.

The development of ANNs dates back to the 1940s [47], but it was not until modern times that they gained widespread acceptance due to advances in computing power. ANNs are organized in layers, with each layer consisting of multiple artificial neurons or units that are connected to neurons in the next layer. The behavior of the network is determined by the activation functions of the individual neurons, which enable the network to detect patterns and learn from previous experiences [47]. Multilayer Perceptron (MLP) is a type of fully connected network that is widely used in ANNs [35], and deep ANNs are those that have two or more hidden layers. An illustration of this is shown in Fig. 2.7.



**Fig. 2.7:** Demonstration of a simple, fully connected artificial neural network.

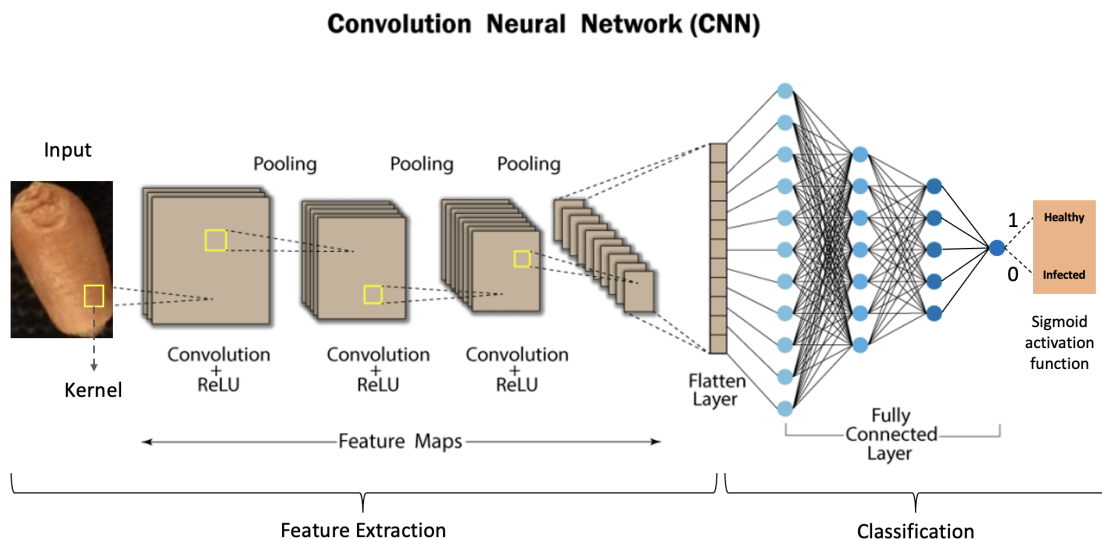
ANNs have revolutionized the field of ML, offering a powerful approach to solving complex classi-

fication and regression tasks [48]. However, their effectiveness is highly dependent on the availability of high-quality training data and careful consideration of the network architecture and activation functions. Ongoing research is focused on developing new architectures and techniques for training ANNs to further enhance their performance in a wide range of applications [49, 50, 51].

### 2.3.4 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a type of ANN that have gained popularity in the field of computer vision due to their ability to extract important features from images. Similar to traditional neural networks, CNNs are composed of multiple stacks of neurons, each with biases and learnable weights [52]. However, in a CNN, only a small region of input layer neurons is connected to neurons in the hidden layer, which are referred to as local receptive fields [53]. These local receptive fields are translated across an image to create a feature map from the input layer to the hidden layer neurons.

One notable aspect of CNNs is parameter sharing, where the weights and bias values are the same for all hidden neurons in a given layer [35]. This means that all hidden neurons are detecting the same feature, such as an edge or a blob, in different regions of the image, making the network tolerant to translation of objects in an image. CNNs are typically composed of several layers [54], including a convolution layer, activation function, pooling layer, and Fully Connected (FC) layer, as shown in Fig. 2.8. The convolution layer extracts main features through operations using a set of filters, while the activation function applies a non-linear transformation to the output of each neuron. The pooling layer reduces the spatial dimensionality of the feature maps, and the FC layer combines the extracted features from the previous layer through a set of linear transformations.



**Fig. 2.8:** Architecture of the CNN used for binary classification of wheat grains. The figure shows the feature extraction process through convolutional layers, followed by classification using fully connected layers with a sigmoid output. The figure is inspired by Developers Breach [link to website].

However, Deep Neural Networks (DNNs) are prone to overfitting due to their large number of parameters, which can lead to poor generalization performance. To address this issue, two popular techniques are used in practice: Dropout and Batch Normalization [55]. Dropout is a regularization technique that randomly drops out a certain percentage of neurons during training to prevent overfitting. This forces the network to learn more robust features and reduces its reliance on any one feature. Batch normalization is a technique for improving the training of DNNs. It normalizes the output of the previous layer by subtracting the mean and dividing by the standard deviation of the activations in a mini-batch. This stabilizes the distribution of inputs to each layer and can accelerate training by reducing the internal covariate shift problem [56]. A phenomenon that occurs during the training of DNNs, where the distribution of the input to each layer changes as the network learns, making training more difficult as each layer must constantly adapt to the changing distribution of its input.

The behavior of the convolution layer can be controlled through several hyperparameters [54], including kernel size, stride, padding, and number of filters. Kernel size determines the size of the sliding window used in convolution, while stride determines how many pixels the kernel window will slide at each step of convolution. Padding determines the amount of pixels or values to include around the border of the image, and the number of filters controls the number of patterns or features that a convolution layer will look for. By adjusting these hyperparameters, CNNs can be optimized for different tasks such as object recognition, object segmentation, detection, and computer vision.

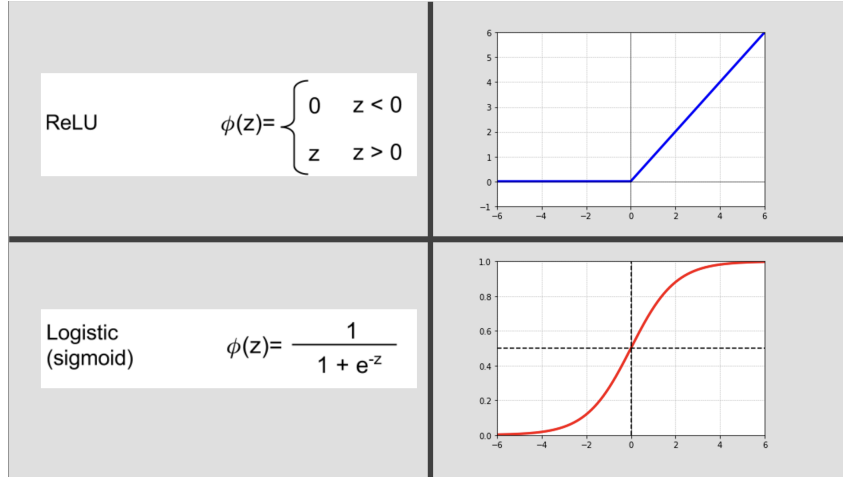
### **2.3.5 Activation and Loss Functions**

The choice of activation functions in a neural network can greatly impact its ability to converge during training, as well as the speed at which convergence occurs [57]. Additionally, activation functions serve to normalize the output of the network to a restricted range, typically  $[0, 1]$  or  $[-1, 1]$ . This is graphically demonstrated in Fig. 2.9.

Sigmoid and Rectified Linear Unit (ReLU) are two widely employed activation functions in the domain of neural networks. Sigmoid is a non-linear activation function that transforms real-valued inputs to a value ranging between 0 and 1. This activation function finds common application in the output layer of binary classification problems, where the objective is to categorize the input data into one of the two classes [35].

In contrast, ReLU is a non-linear activation function that takes negative input values and maps them to 0, and positive input values to themselves. The simplicity and computational efficiency of ReLU makes it a popular choice for use in hidden layers of DNNs [58].





**Fig. 2.9:** Comparison of the Sigmoid and ReLU activation functions commonly used in neural networks. The Sigmoid function maps inputs to a smooth output range between 0 and 1, while the ReLU function maps inputs linearly to a range between 0 and positive infinity.

Effective activation functions must meet several criteria, including differentiability over the entire domain, the inclusion of nonlinearities, computational efficiency, and the ability to prevent the exploding or vanishing gradient problem [59].

The vanishing gradient problem arises when the gradients of the loss function with respect to the parameters of the network become extremely small as they propagate through the layers of the network [35]. This problem is particularly prevalent in DNNs with many layers, especially those using activation functions with derivatives that approach zero as the input becomes very large or very small. Consequently, the weights of the network may be updated very slowly or not at all in some layers, resulting in poor convergence or stagnation of the training process which in some cases can trap the model in a local minimum [35, 60], preventing it from reaching the global minimum of the loss function.

In the training of neural networks, the choice of loss function is also crucial. Working with binary classification, showed in Eq. (3), binary cross-entropy is applied. The loss function is used to evaluate the difference between the predicted output of the network and the true output, and the goal of training is to minimize this difference [61].

$$H(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i) \quad (3)$$

where  $H(q)$  refers to the entropy, the measure of the uncertainty associated with a given distribution  $q(y)$ . Meanwhile,  $y$  is the true label (either 0 or 1) and  $\hat{y}$  is the predicted probability of the positive class [62].  $N$  corresponds to the number of training samples.



### 2.3.6 Metrics for Binary Classification

In order to evaluate the performance of a model, various metrics may be utilized. A confusion matrix serves to depict the model's ability to correctly and falsely predict outcomes. Fig. 2.10 provides an illustrative example of such a matrix, which accounts for True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN) values. The values within the confusion matrix may further be employed to calculate additional metrics such as precision, recall, and F-score.

		Predicted Label	
		P	N
True Label	P	TP True Positive	FN False Negative
	N	FP False Positive	TN True Negative

**Fig. 2.10:** Example of a confusion matrix for binary classifiers.

Precision and recall represent two distinct measures, wherein precision is particularly relevant when minimizing FPs is of primary concern, and recall is a suitable metric when the focus is on minimizing FNs [63]. Precision and recall may be calculated using Eqs. (4 & 5), respectively [35]. Often, the performance evaluation of classification problems requires the consideration of both precision and recall metrics, since a good recall score may not guarantee a good precision score, and vice versa. In response to this limitation, the F-score (6), also known as the F1-score, has been introduced as a composite measure to express both precision and recall concerns through a single score. The F-score is a prevalent metric utilized in the assessment of binary (or multiclass) classification problems [35].

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{FN + TP} \quad (5)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

## 2.4 Breeze Software

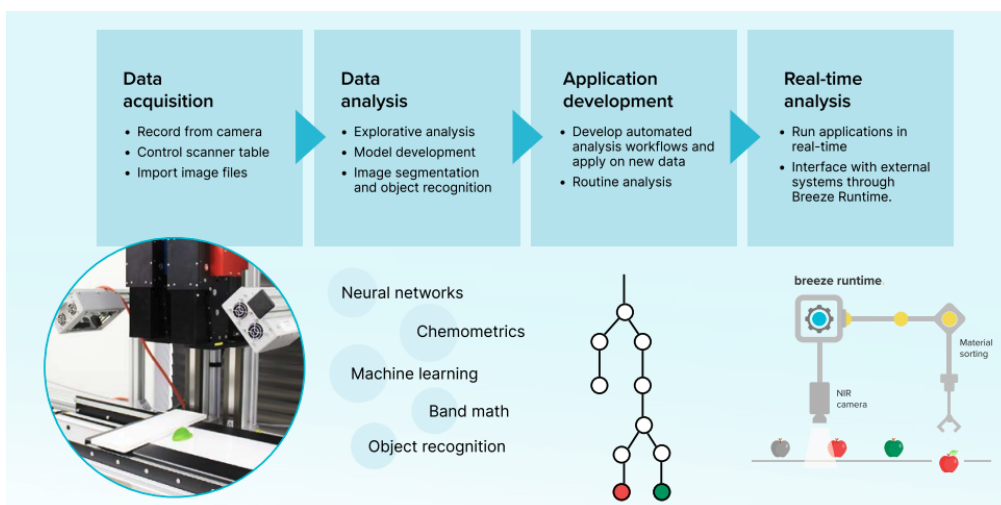
Breeze is a HSI software developed by Prediktera AB. It enables a wide range of HSI applications, used for research, application development, routine analysis and also real-time industrial solutions [3]. Prediktera was established as a result of research in multivariate data analysis conducted at Umeå University in Sweden. Specializing in HSI software, the company has dedicated its efforts towards this field since 2015. Recently, in 2022, Prediktera was acquired by Norsk Elektro Optikk AS (NEO), a renowned electro optics enterprise in Scandinavia, and now operates as a wholly owned subsidiary of NEO [3].

The Breeze software follows a systematic workflow consisting of several steps for processing and analyzing hyperspectral images. The workflow commences with the importation of data, which includes hyperspectral images and any relevant metadata. Subsequently, the data undergoes pre-processing, where procedures such as noisy pixel removal, normalization, and baseline correction are applied to improve the quality of the data.

Next, Breeze utilizes image segmentation to partition hyperspectral images into smaller regions based on their spectral similarity. Following segmentation, features are extracted from each region, which represent the characteristics of the corresponding segment. The software then employs modelling techniques, such as PLS regression, PCA, and DA, to develop predictive models for classification, regression and prediction of new samples.

The developed models undergo validation using techniques such as cross-validation to evaluate their predictive capabilities. Finally, the models are applied to new samples or images to make predictions or classifications, and the output can be visualized using different tools such as classification tables and scatter plots.

The steps describing how the software functionate is shown in Fig. 2.11.



**Fig. 2.11:** General workflow of Breeze software from Prediktera AB [3].

As previously mentioned, the software exhibits versatility and can be applied in various applications, including agriculture. Its utilization in grain analysis is particularly beneficial. To this end, a

study published by Kate Sendin, et al. from Department of Food Science, Stellenboshch University, South Africa [64] employed the software to distinguish sound from infected maize kernels. The researchers segmented a total of 2088 samples and calculated the average spectra for each seed. The dataset were processed using the software to identify unwanted pixels such as background, outliers, and edge effects. Subsequently, a PLS-DA model was built, which accurately classified each object as either sound or infected with the fungus, fusarium. The model underwent validation, and upon application to several recorded images, achieved an accuracy of 93.3%. This output model could serve as a tool for conducting real-time quality control and detecting defective kernels.

## 2.5 Aquaphotomics

Aquaphotomics is an approach to analyze molecules based on their absorbance spectra in the NIR range [65]. The study of water absorbance bands and absorbance patterns can offer valuable insights into the structures of water and its inherent interactions with other matters of the aqueous system [66]. The underlying principle of this concept is that the absorbance spectrum related to water can be analyzed through its interaction with light, thus providing a rich source of physical and chemical information [67]. Water molecules, owing to their hydrogen bonding, can assemble into a layer that surrounds a solute, forming a hydration shell. The presence of a hydration shell can significantly influence the ability to maintain information about the spectra and disturbances related to it. The *OH* bonds absorption region of water that corresponds to the first overtone of *OH*, spans in the range of 1300-1600 *nm* and are often referred to as Water Matrix Coordinates (WAMACS) [68]. The characteristic spectral pattern and composition of the WAMACS are collectively referred to as Water Spectral Pattern (WASP). The WAMACS encompass 12 distinct bands, as illustrated in Table 2.1. To analyze the behavior of individual WAMACS bands and their interactions, a normalized absorption aquagram can be plotted to compare the absorbance values at different WAMACS positions between healthy and infected grains.

$$A'_\lambda = \frac{A_\lambda - \mu_\lambda}{\sigma_\lambda} \quad (7)$$

where  $A'_\lambda$  corresponds to the normalized absorbance spectra showcased on the aquagram. Meanwhile,  $A_\lambda$  is the absorbance after EMSC.  $\mu_\lambda$  is the mean, and  $\sigma_\lambda$  is the standard deviation, of all spectra for the examined group of samples after correction.  $\lambda$  is the wavelengths related to the chosen WAMACS.

**Table 2.1:** The table presents the ranges of WAMACS in the first overtone, their corresponding wavelengths, descriptions and assignment & vibration types. Including  $S_{0-4}$ , hydration shell, stretching and bending.

<b>WAMACS</b>	<b>Wavelengths [nm]</b>	<b>Description</b>	<b>Assignment &amp; Vibration</b>
C1	1336–1348	$H_2O$ asymmetric stretching vibration	$\nu_3$ & Asymmetric Stretching
C2	1360–1366	Water solvation shell, $OH-(H_2O)_{1,2,4}$	Hydration shell
C3	1370–1376	Symmetrical stretching fundamental vibration and $H_2O$ asymmetric stretching vibration	$\nu_1 + \nu_3$ & Symmetric- and Asymmetric Stretching
C4	1380–1388	Water solvation shell, $OH-(H_2O)_{1,4}$ and superoxide, $O_2 - (H_2O)_4$	Hydration shell
C5	1398–1418	Free water and free $OH-$	S0
C6	1421–1430	$H - OH$ bend and $O...O$	Bending
C7	1432–1444	One hydrogen-bonded molecule	S1
C8	1448–1454	Water solvation shell, $OH - (H_2O)_{4,5}$	Hydration shell
C9	1458–1468	Two hydrogen-bonded molecules	S2
C10	1472–1482	Three hydrogen-bonded molecules	S3
C11	1482–1495	Four hydrogen-bonded molecules	S4
C12	1506–1516	Symmetrical stretching fundamental vibration, and doubly degenerate bending fundamental	$\nu_1, \nu_2$ & Symmetric Stretching and Bending

### 3 Methodology

The experiments practiced through the thesis will be disclosed in this section. To avoid inhaling dangerous mycotoxins from the wheat grains, certain precautions were taken into consideration. Gloves, glasses and face masks were worn, along with an exhaust fan, through the entire lab work process. Further, all data were analyzed using Breeze, in addition to Python 3.9.7 along with the mentioned python packages.

All kernels included in the images were manually annotated before being captured. However, manual annotation is not without error due to various factors, such as fatigue, external distractions, perceptual biases, and inconsistencies while assigning annotations [69]. Therefore, it is crucial to understand that this reference is not an absolute ground truth and that this fact must be taken into account when interpreting the results. While the ground truth provides a valuable point of comparison, it is subject to limitations and potential errors that may affect its accuracy. Therefore, it is essential to approach the results with a critical eye and take into account the possibility of discrepancies or uncertainties.

#### 3.1 Data Collection

##### 3.1.1 Grains

Wheat grains were gathered from Vollebekk experimental farm, located close to NMBU campus in Ås municipality, Norway. In total, 20 different varieties of wheat grains was sorted and analyzed. For each variety, shown in Table 3.1, a manual annotation of 20 healthy and 20 infected seeds was done. This gives a total of 800 seeds in total as the foundation dataset. The grains were randomly picked out from the blue bags shown in Fig. 3.1.



**Fig. 3.1:** Preparation and annotation of samples in the NMBU workshop.

In order to distinguish healthy and infected grains, the chosen samples were distributed by visual assessment at the NMBU workshop on 19<sup>th</sup> and 20<sup>th</sup> of January. Following the annotation, each grain was carefully placed in an individual container equipped with a lid and a label indicating the corresponding route, as well as denoting the grain's condition as either "H" for healthy or "I" for infected.

**Table 3.1:** Overview of the 20 different wheat varieties and their concentration of Fusarium Damaged Kernels (FDK) in percentage and deoxynivalenol (DON) contamination.

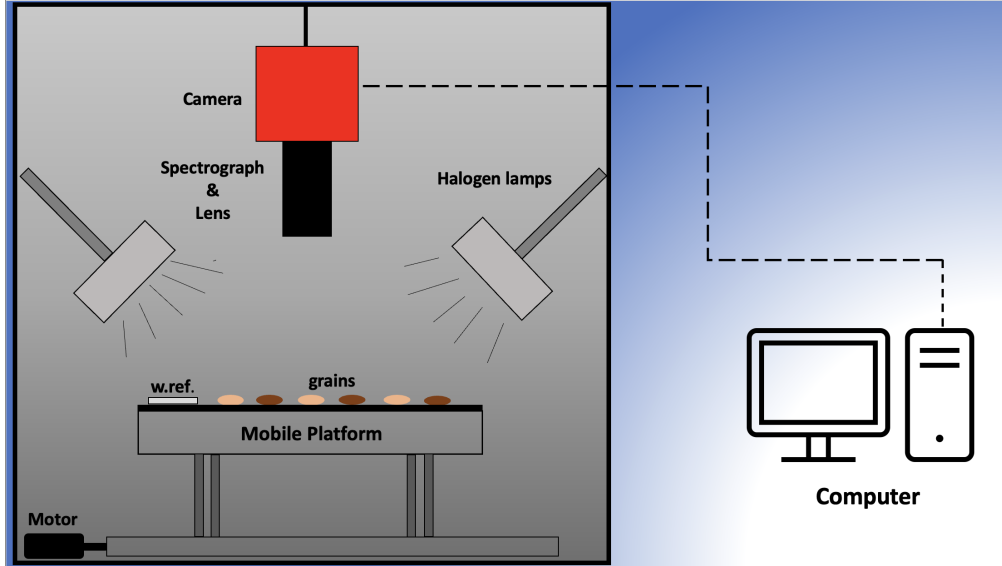
Route	Name	FDK (%)	DON
4205	Saar	30	33.4
4208	N894037	2	10.5
4217	Vinjett 5 (susceptible)	45	34.6
4222	Zebra	20	27.2
4223	MILAN/SHA7	30	40.2
4226	BCN*2//CROC_1/AE,SQUARROSA (886)	10	20.3
4228	Krabat	40	23.6
4234	Sumai #3-1 (12SRSN)	1	13.2
4302	Bjarne	30	22.3
4303	Demonstrant	40	24.5
4310	EMB16/CBRD//CBRD	40	32.9
4319	SABUF <sup>1</sup>	25	28.9
4339	Seniorita	30	16.8
4340	GUAM92//PSN/BOW	40	55.1
4423	GONDO	40	61.1
4428	CJ9403	45	31.3
4437	Mirakel	10	14.2
4505	Arabella	25	15.3
4525	Bastian	18	24.2
4532	NG8675/CBRD//SHA5/WEAVER	15	30.3

### 3.1.2 Image Acquisition

The imaging system used to capture the images analyzed in this thesis is HySpex SWIR-384 camera, which has a wavelength range of approximately 950-2500 *nm*, divided into 288 bands. The pushbroom scanning camera also had a frame period of 5000 *ms* and an integration time of 4800 *ms*. The hyperspectral camera setup comprises of several components, including a lens, a plate, two halogen lamps powered by a 12V *DC* power supply, custom designed to focus the illumination into a line that overlaps with the camera's FOV, a mobile platform that can be adjusted to various heights and a white reference plate (60% reflectance). The plate serves as a stable platform for the objects being imaged, while the halogen lamps, which emits a continuous spectrum of light that

<sup>1</sup>Full name of the wheat variety: SABUF/5/BCN/4/RABI//GS/CRA/3/AE,SQUARROSA (190)

covers a broad range of wavelengths, provide illumination to ensure clear and detailed imaging. The mobile platform allows for the camera to be positioned at different heights to capture images of grains at a certain focus point displayed at the monitor. This was a crucial process to get the best possible resolution for the images. Finally, the white reference plate is used to calibrate the hyperspectral images. The camera setup is illustrated in Fig. 3.2.



**Fig. 3.2:** Setup for image acquisition in laboratory.

Initially, the grains were placed in a petri dish and captured using a 30cm lens for the hyperspectral camera on the 25<sup>th</sup> of January. However, the analysis of the images in Python revealed that the shadows from the grains presented a challenge due to the white background. Therefore, the next two days, 26<sup>th</sup> and 27<sup>th</sup> of January, new images were taken with a black fabric as the background beneath the petri dish. Later, 16<sup>th</sup> and 17<sup>th</sup> of February, subsequent to the acquisition of these initial images, a new macro lens became available and was used in the imaging process, which yielded images with increased details and was deemed to be a superior approach for grain analysis for this thesis. As a consequence, utilizing the macro lens led to a narrower FOV. In comparison with the 30cm lens which has a FOV  $\sim 9.5cm$ , the macro lens displays a reduced FOV  $\sim 2cm$ . The kernels were aligned in two or three rows, with 20 kernels per image, to assure coverage within the FOV. To ensure accurate alignment, the wheat kernels were manually positioned on a black fabric using tweezers. This was done to ensure that the kernels were adequately spaced apart, as clustering could lead to errors in the algorithm treating multiple kernels as a single entity during further analysis. Additionally, to introduce variability in their orientations, the kernels were placed randomly before the image scan.

RGB images were acquired using a Nikon Df camera with a color profile of sRGB IEC61966-2.1, which is a standard color space used in digital imaging. The camera's optical system had an f-ratio of 6.3, representing the ratio of the system's focal length to the diameter of the aperture. A larger f-value indicates a smaller aperture diameter, which limits the amount of light entering the camera and results in a wider FOV. In contrast, a smaller f-value, as used in this thesis, implies a larger aperture diameter that allows more light to enter the camera and results in higher image resolution and finer details. The resolution of the images based on these specifications was 300 pixels per

inch, indicating the level of details captured within each inch of the image. Directly after the scan of the hyperspectral image, the RGB image of the same route was acquired with Nikon camera by covering the front of the halogen lamps with white A4 paper in order to reduce specular reflection. At first, along with the 30cm lens, RGB images were acquired with a white background and light source corresponding to natural sunlight. These images were not of the best resolution. Therefore, new RGB images were also taken alongside the makro hyperspectral images, now with a black background and light from the halogen lamps. This approach provided a closer look, elimination of shadows and more nuanced light, enabling greater detail to be captured in the images.

### 3.2 Calibration and Preprocessing Techniques

All the hyperspectral images, both taken with the 30cm- and macro lens were post-processed in order to mitigate from external interference. This was done by carrying out a white- and dark current correction. Background radiation may disturb the sensors in the camera and thereby give unwanted results. Therefore, applying dark current correction for each image is desirable. This was done by using a software called RadV2 provided by the producer of the hyperspectral camera [Hypex.com].

$$I_{Calibrated} = I_{Raw} - I_D \quad (8)$$

where  $I_{Calibrated}$  is the calibrated image for the dark current,  $I_{Raw}$  is the original raw hyperspectral image and  $I_D$  is the dark current image.

In addition, to correct for the differences in light source, it is possible to calibrate the image with Eq. (9)

$$I_{Calibrated} = \frac{I_{Raw}}{I_W} \quad (9)$$

where  $I_{Calibrated}$  is the corrected image,  $I_{Raw}$  is the original image, while  $I_W$  is considered as the white reference containing the mean values for each column in each band in the image.

When combining Eq. (8) with Eq. (9), we can obtain the following

$$I_{Calibrated} = \frac{I_{Raw} - I_D}{I_W - I_D} \quad (10)$$

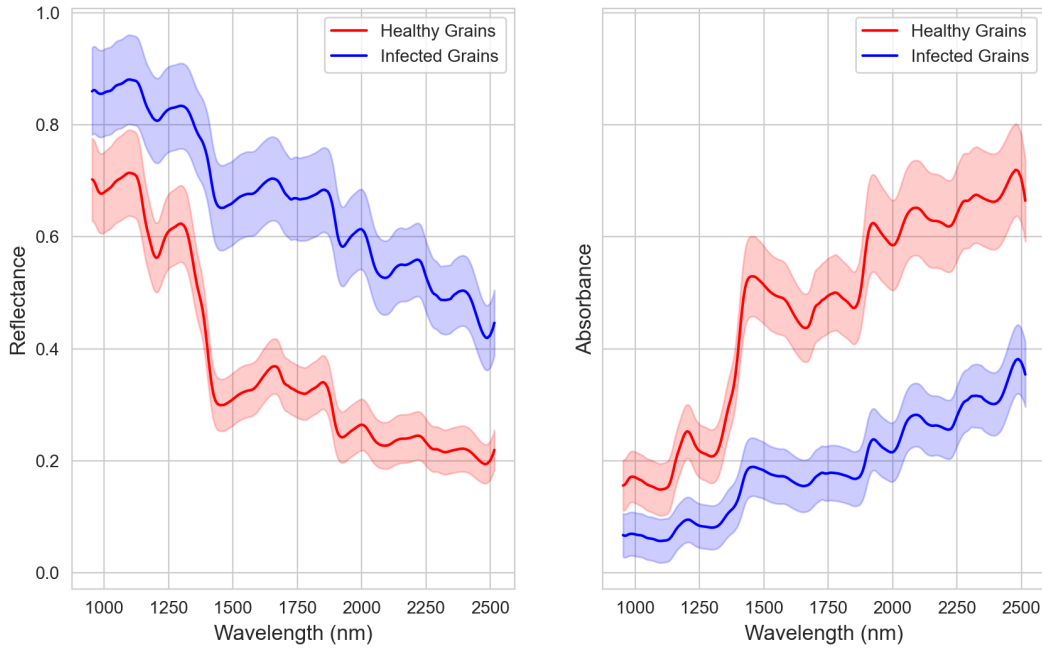
where  $I_{Calibrated}$  now corresponds to the normalized raw hyperspectral image.



When analyzing properties of grains, such as in aquaphotomic analysis, there is often a requirement to transition from reflectance- to absorbance spectrum. This transition facilitates the examination of various characteristics and can be accomplished by using Eq. (11).

$$A = \log_{10} \left( \frac{1}{R} \right) \quad (11)$$

where  $A$  is the absorbance and  $R$  is the reflectance. The reflectance and absorbance spectra for healthy and infected grains after calibration is shown in Fig. 3.3.



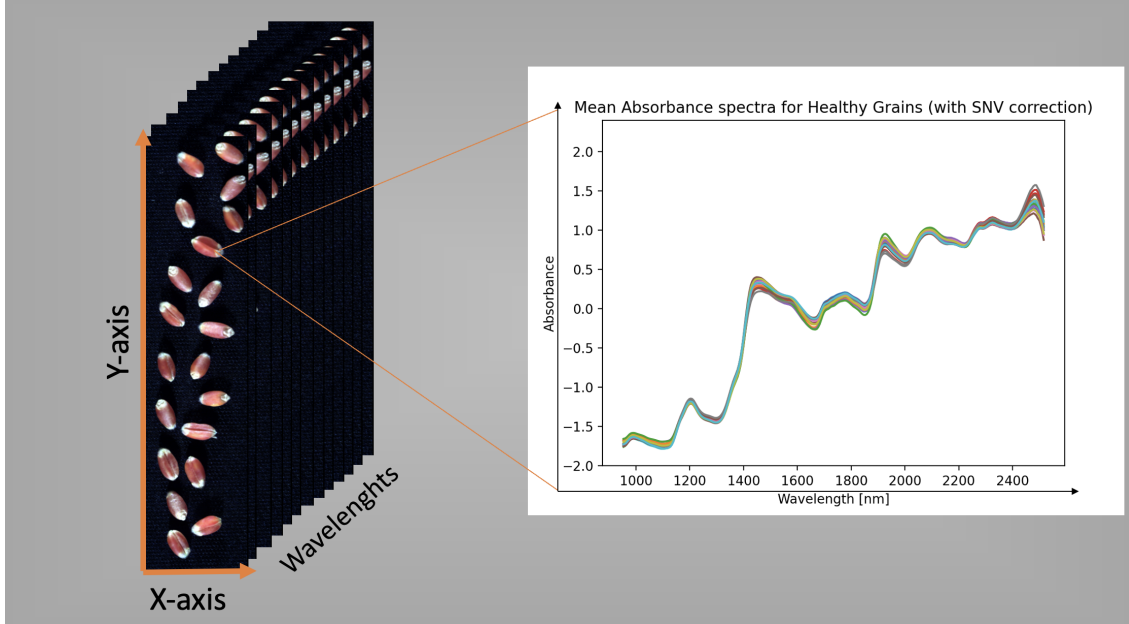
**Fig. 3.3:** Mean spectra with shaded areas representing the standard deviation for reflectance and absorbance of all healthy and infected grains in the dataset.

Prior to analyzing the images in breeze and python, various preprocessing techniques were applied to the images in order to reduce noises and disturbances, with the goal of ensuring accurate and reliable spectral measurements.

Standard Normal Variate (SNV) is a preprocessing technique widely used in spectroscopy to correct for multiplicative variations between spectra that can arise due to differences in sample path length or physical properties, sample preparation, and spectrometer optics [70]. These variations can confound with the signal from changes in component concentrations, making quantitative applications difficult.

SNV correction involves subtracting the mean intensity from each variable intensity in a spectrum and dividing the resulting values by the standard deviation, so that the resulting intensities have a mean of zero and a standard deviation of one. This normalization process is performed on a single spectrum, and no reference spectrum is required for the correction [71]. Different surface

roughness, incident angle of the light and distinct geometries can greatly affect the type of scattering [72]. Hence, SNV correction (Eq. 12) is particularly useful to minimize scattering effects when working with uneven objects like grains. Fig. 3.4 depicts a hyperspectral image of healthy grains and its corresponding SNV-corrected (absorbance) spectrum.



**Fig. 3.4:** Hyperspectral image of healthy grains. Showing a 2-dimensional image with the corresponding (SNV corrected) mean absorbance spectra for each of the 20 grains in the hyperspectral image.

$$X_i^{SNV} = \frac{X_i - \bar{X}_i}{\sigma_i} \quad (12)$$

where  $X_i^{SNV}$  is the corrected spectrum for each pixel  $i$ ,  $X_i$  is the original spectrum, while  $\bar{X}_i$  and  $\sigma_i$  corresponds to the mean and standard deviation, respectively.

Another method which is similar to SNV is Multiplicative Signal Correction (MSC). This method also aims to reduce scattering effects based on particle distribution and particle size [73]. A more advanced version of MSC, known as Extended Multiplicative Signal Correction (EMSC), is based on the fundamental principle of Lambert-Beer's law [74]. The approach involves fitting a regression model to each spectrum in an image based on a reference spectrum, and subsequently using the derived coefficients to perform the necessary correction.

The absorbance spectrum  $A$  of a given pixel  $i$  can be represented mathematically using the Eq. (13).

$$A_i = a_i + b_{1,i} * \bar{x} + b_{2,i} * w + b_{3,i} * w^2 + E_i \quad (13)$$

This equation incorporates the EMSC approach, which is distinct from the commonly used MSC method, as it includes the effects of both linear and quadratic baseline variations ( $w$  and  $w^2$ , respectively) [75]. Additionally, the equation includes an additive baseline variation term ( $a_i$ ) and scaling parameters ( $b_{i,1,2,3}$ ) that approximate the spectrum through least square estimation. Meanwhile,  $\bar{x}$  is the reference spectrum which is usually gathered by taking the mean of all spectrum. The term  $E_i$  accounts for unmodeled variations that arise from both chemical variations and measurement noise.

The corrected spectra can be calculated by using Eq. (14)

$$A_{i,corr} = \frac{A_i - a_i - b_{2,i} * w - b_{3,i} * w^2}{b_{1,i}} \quad (14)$$

One important take of EMSC is that the approach allows for the correction of baseline variations, which are often a confounding factor in spectroscopic analyses, and improves the accuracy of the measured spectra [74].

Hyperspectral image data acquired from grains can also be analyzed through the use of first and second derivative spectra, to identify and investigate specific spectral features of interest [76]. The first derivative spectrum can be calculated using Eq. (15)

$$\frac{dy}{d\lambda} = \frac{y_{i+1} - y_{i-1}}{2\Delta\lambda} \quad (15)$$

where  $y_i$  denotes the spectral intensity at wavelength  $\lambda_i$ ,  $\Delta\lambda$  represents the spacing between neighboring wavelengths, and the derivative  $\frac{dy}{d\lambda}$  is evaluated at the central wavelength  $\lambda_0$ . By taking the difference between neighboring data points and dividing by the spacing between the data points, the resulting first derivative spectrum highlights the locations of zero-crossings in the data, which can be indicative of the locations of peaks and valleys in the original spectrum. [77].

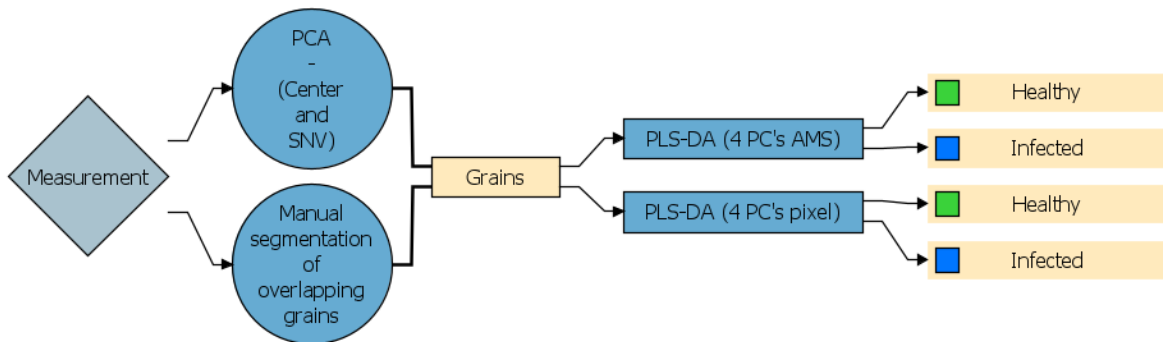
To compute the second derivative spectrum, the difference between the neighboring points of the first derivative spectrum is taken and divided by the spacing between the data points. Eq. (16) represents the calculation of the second derivative spectrum  $\frac{d^2y}{d\lambda^2}$  taken at the central wavelength  $\lambda_0$ .

$$\frac{d^2y}{d\lambda^2} = \frac{y_{i+1} - 2y_0 + y_{i-1}}{\Delta\lambda^2} \quad (16)$$

The resulting second derivative spectrum highlights inflection points in the data [77], which can be instrumental in identifying subtle spectral features that may not be apparent in the raw data or first derivative spectrum.

### 3.3 Breeze Workflow for Spectral Imaging

Before doing any computations with breeze, the raw hyperspectral images were imported into the software along with their corresponding HDR files, which contain important metadata about the images. Once the images were imported, they were then categorized into two groups, healthy and infected, to be used as labeled data in later classification tasks. After this, a PCA model was established to segment out the grains for further analysis. The workflow process is illustrated in Fig. 3.5.



**Fig. 3.5:** Analysis tree of the workflow in Breeze. From segmentation to classification. The first PLS-DA model contains four PC's based on an object-based evaluation, whereas the other PLS-DA model contains four PC's based on a pixel-based classification.

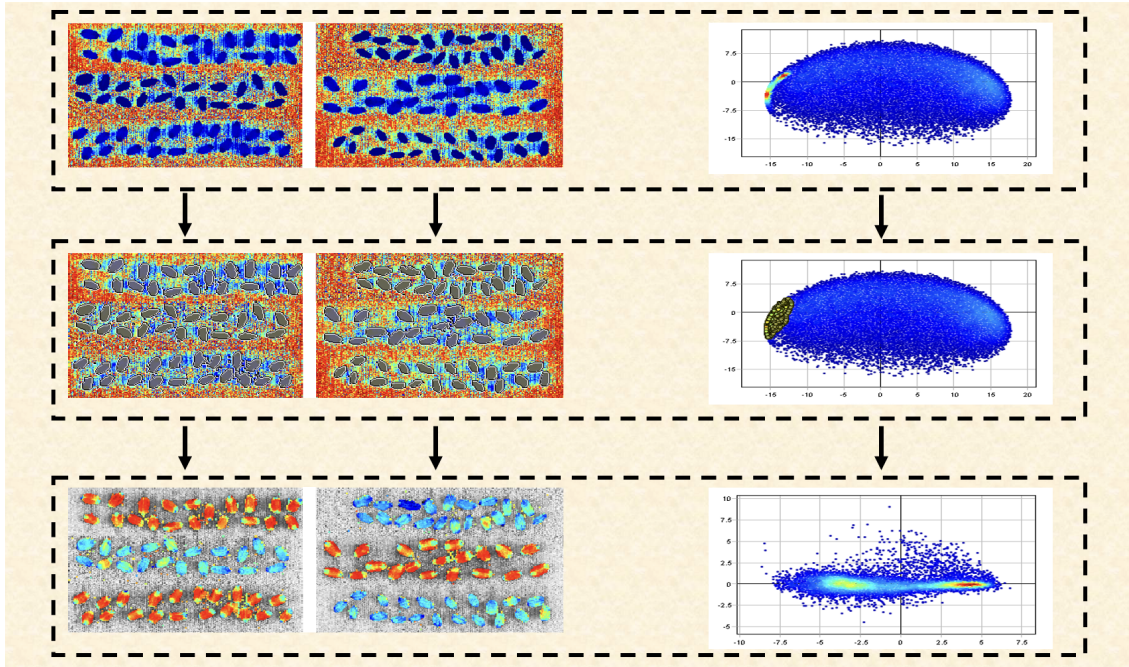
#### 3.3.1 PCA - Feature Extraction and Segmentation

Six representative images were selected from the dataset to use as input for segmentation. In order to minimize noise and disturbances, specific wavelengths were chosen and pretreatments were applied. To begin with, the 10 first and last bands were dropped in order to reduce noise. Consequently, SNV pretreatment was mostly utilized, but later on, different pretreatments such as first- and second derivative was tested as well.

After preprocessing, cluster regions were cropped out by marking an eclipse around the Region of Interest (ROI) in the scatter plot. This was done to remove all background pixels and only include pixels related to grains. The process is shown in Fig. 3.6.

Next, a critical distance model was used to segment the remaining pixels based on their similarity or dissimilarity to neighboring pixels. The critical distance was adjusted by dragging the Dcrit line horizontally to include additional pixels that may have been missed in the initial segmentation. However, this approach occasionally resulted in grouping multiple grains together, requiring manual segmentation to separate them later on.

In the end, a minimum area size threshold of 1000 pixels was applied to the segmented regions, filtering out any regions that were smaller than the defined threshold. This ensured that only large enough regions, corresponding to individual grains, were included in the analysis.



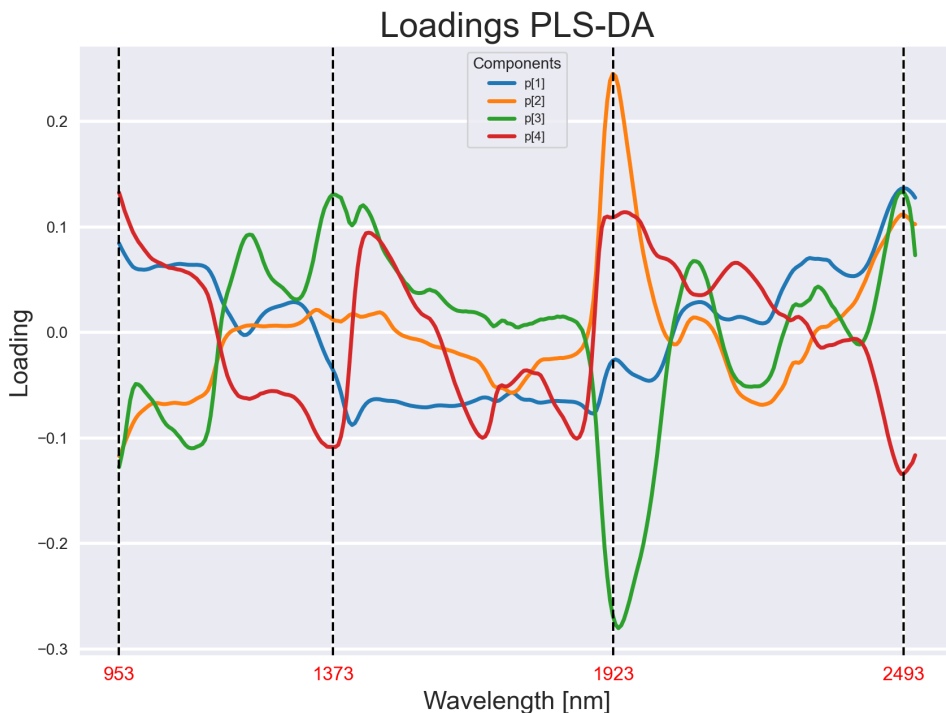
**Fig. 3.6:** Process of pixel segmentation based on a scatter plot. After selecting a wavelength interval and applying pretreatment, the top image displays six raw images including background pixels, while the middle image shows the ROI. The last image shows the six resulting images after background removal, along with their correlated scatter plot.

### 3.3.2 PLS-DA - Classification

Two distinct methods were employed for classification in this study, including a pixel-based approach using Pixel Majority (PM), and an object-based approach using Average Mean Spectrum (AMS). The PM approach involved classification of individual pixels based on the majority of its neighboring pixels, while the AMS approach classifies each object based on its average spectrum. For both approaches, the final PLS-DA model attained for this study consisted of using four principal components with SNV correction. As previously mentioned, first- and second-derivative pre-treatments were also applied to the data, but no improvement in classification results was observed. Therefore, SNV was ultimately selected as the preferred pretreatment method. The model was trained on a total of 640 samples, while 160 samples were assigned for testing purposes. Initially, all bands were included in the model, given the high importance value of the first and last bands. Later, the model was also tested to determine if it could achieve similar results using only a few effective wavelengths.

### 3.3.3 Effective Wavelengths

Based on a loading plot from the PLS-DA model, four wavelength seemed to be of significant importance when doing prediction with the model. As Fig. 3.7 shows, these wavelength were 953 nm, 1373 nm, 1923 nm and 2493 nm. After completing the training, these loadings were exported from the model and imported into a dataframe in python for visualization and interpretation. It was observed that these four wavelengths exhibited a strong negative correlation between their components. By using only these four wavelengths, the model's efficiency and cost-effectiveness could potentially be improved significantly.



**Fig. 3.7:** Loading plot from the PLS-DA model showing the corresponding wavelengths and their importance for binary classification of the grains. The four specific wavelengths highlighted in red indicates a strong negative correlation between the two classes.

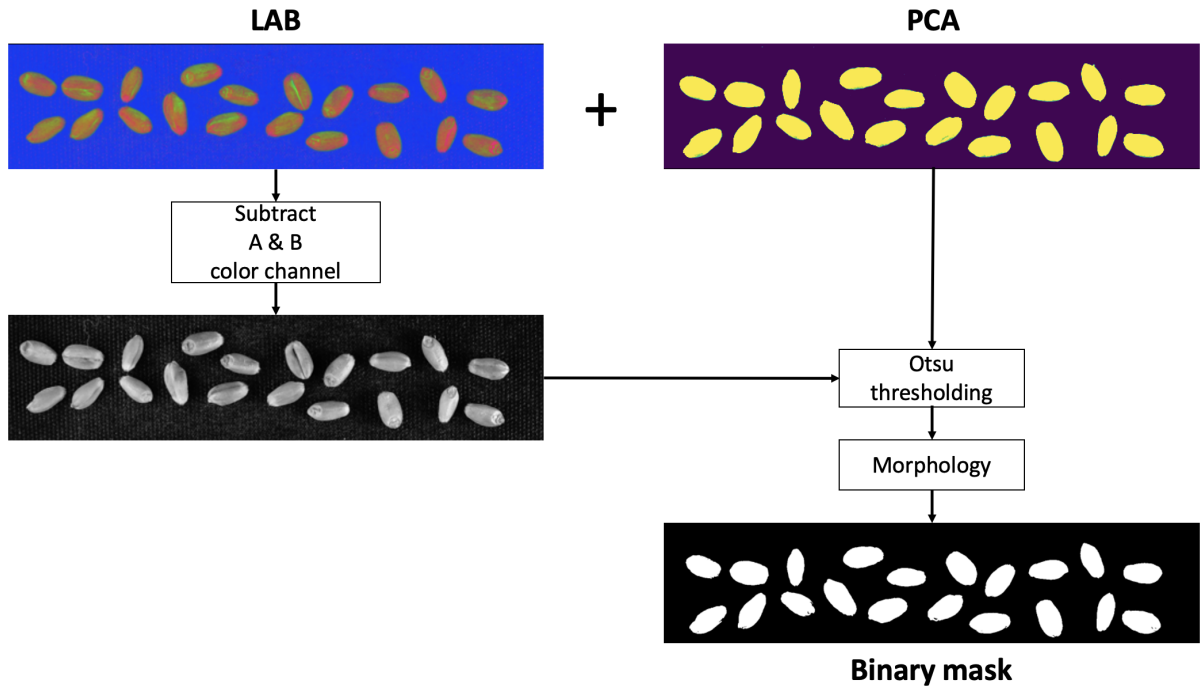
## 3.4 Predicting Grain Condition using CNN with RGB Images

### 3.4.1 Mask Generation and Preparation of Dataset

Based on the findings from the effective wavelengths using hyperspectral images, it became interesting to investigate if comparable results could be achieved using RGB images. Thus, a simple CNN was developed to analyze individual grains, requiring the creation of a mask for each grain in an image using the "rgb2lab" package from the scikit-image library in python. Although both color information channels (a and b) were attempted for this purpose, the resulting mask still had some undefined spots. However, the L-channel, which represents the image's lightness, was found

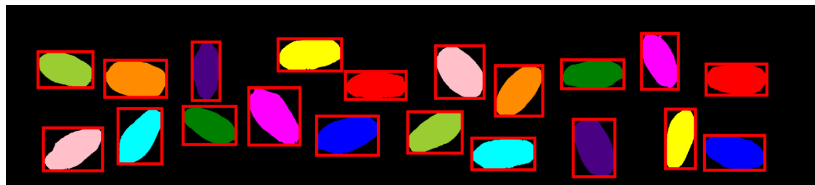


to be more suitable for mask creation, and applying Otsu thresholding to it improved the mask quality. Another effective approach for generating the mask was to use PCA based on the spectral python package and retaining 99% of the image's variance. To further eliminate any remaining noise in the image, the morphology was used, employing the "remove\_small\_objects" and "remove\_small\_holes" packages to avoid the misidentification of small objects as grains. The final binary mask was then generated, as illustrated in Fig. 3.8.



**Fig. 3.8:** Process of generating binary mask using a combination of LAB color space and PCA, followed by thresholding and morphology operations.

To extract the bounding box of the individual grains from the binary mask, each grain were assigned a unique label through the use of "label" and "regionprops" packages from the scikit-image library, as shown in Fig. 3.9. These bounding boxes were subsequently utilized to crop out and store the 800 grains as data images for the CNN model.



**Fig. 3.9:** Mask for RGB images with one label for each grain color-coded and their corresponding bounding boxes highlighted in red.

### 3.4.2 Parameter Optimization

In order to obtain the most accurate and robust results from a CNN model, hyperparameter optimization is a crucial step. In this thesis, the GridsearchCV method from the scikit-learn library were implemented to perform an exhaustive search over a specified parameter grid. The parameter grid, presented in table 3.2 included various combinations of hyperparameters, such as the number of filters in the convolutional layer, kernel size, number of units in the dense layer, and learning rate.

**Table 3.2:** Table representing the hyperparameters used for optimization when building the CNN, with a total of 96 possible combinations.

Hyperparameter	Values
Filters	[3, 6, 9, 12]
Kernel size	[2, 3, 5]
Dense units	[3, 6, 18, 24]
Learning rate	[0.0001, 0.001]

This approach helped to avoid the tedious and time-consuming process of manual guessing, ultimately leading to a significant improvement in the accuracy and performance of the model.

### 3.4.3 Data Augmentation

To enhance the generalization ability and increase the amount of training data in the study on wheat grains, data augmentation techniques were employed. The ImageDataGenerator class from the Keras library was used to generate new training images by applying various transformations to the existing dataset, including a rotation range of 45 degrees, width and height shift ranges of 0.2, horizontal and vertical flips, zoom range of 0.3, and nearest neighbor fill mode. These transformations introduced variability to the images, making the model more robust and less prone to overfitting. Additionally, the use of data augmentation increased the size of the training dataset, which is particularly useful in cases where the amount of available data is limited. This, in turn, enabled the development of a comprehensive model capable of effectively capturing the features of wheat grains and accurately predicting their condition.

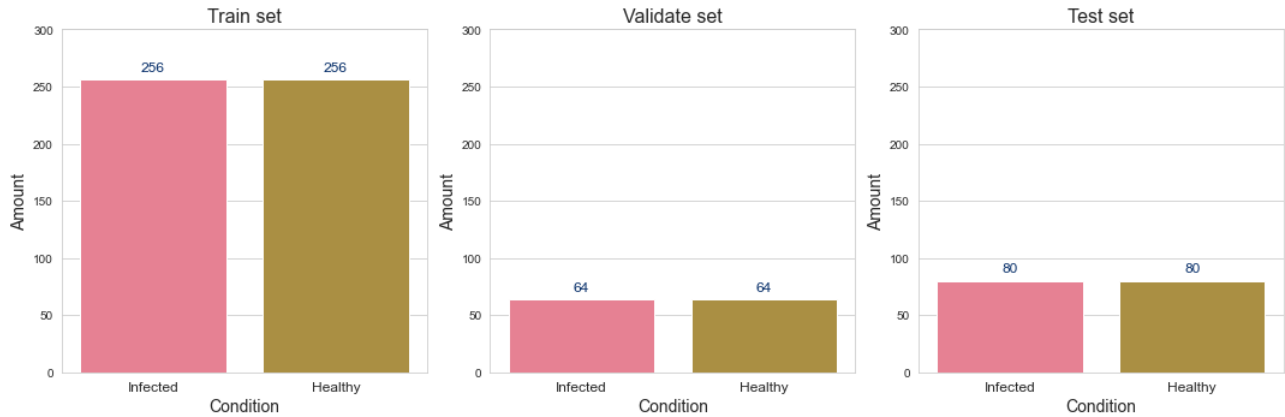
### 3.4.4 Data Distribution

The dataset used in this study was split into train, validation, and test sets, and the pixel values were normalized to the range [0, 1]. To ensure adequate data for model training and validation, the train and validation sets were further divided into smaller sets. The dataset consists of 512 images in the train set, 128 images in the validation set, and 160 images in the test set. Each image has a width and height of 128 pixels and three channels corresponding to the RGB color space. The distribution and shape of the dataset are shown in Fig. 3.10 and Table 3.3, respectively.



**Table 3.3:** Shape of the train-, validation- and test data

	<b>Images</b>	<b>Width</b>	<b>Height</b>	<b>Channels</b>
<b>Train Data</b>	512	128	128	3
<b>Validation Data</b>	128	128	128	3
<b>Test Data</b>	160	128	128	3



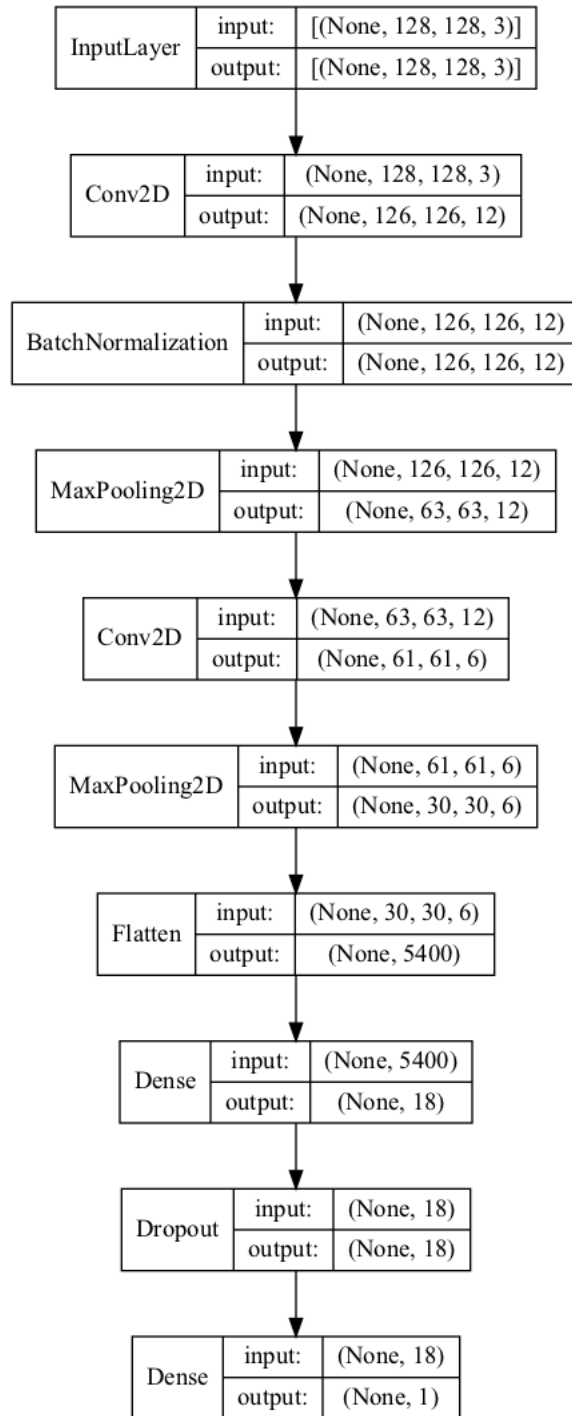
**Fig. 3.10:** Distribution of the training, validation- and test set.

### 3.4.5 CNN Architecture

In order to construct an effective model architecture, a CNN was designed, consisting of convolutional and pooling layers, followed by a flatten layer, two dense layers, and a batch normalization and dropout layer to prevent overfitting. As previously stated, augmented data generated by the ImageDataGenerator function in the Keras library was used to train the model to ensure that the model could generalize well to new data.

Moreover, hyperparameter tuning was performed to identify the optimal model configuration. A grid search was conducted to explore various combinations of hyperparameters, as listed in Table 3.2, with the best performing configuration selected based on cross-validation performance. The optimal hyperparameters were found to be 12 filters in the first convolution layer with a stride of one and valid padding, followed by six filters in the second convolution layer with the same stride and padding configuration. Valid padding means that the convolution operation is applied only to the valid positions of the input, resulting in an output feature map that is smaller than the input. This was used in both convolutional layers, along with a 3x3 kernel size. Additionally, the model included 18 dense units and was trained with a learning rate of 0.001.

The model was compiled using binary cross-entropy loss and Adam optimizer with the optimal learning rate, and subsequently trained with a batch size of 32 over 50 epochs. Performance evaluation was performed using the validation set. The architecture of the model is illustrated in Fig. 3.11.



**Fig. 3.11:** A schematic representation of the CNN architecture used for the binary classification task. The model consists of two convolutional layers with max pooling and batch normalization, followed by a fully connected layer with dropout and a sigmoid output layer.

### 3.5 Exploring Alternative Machine Learning Approaches

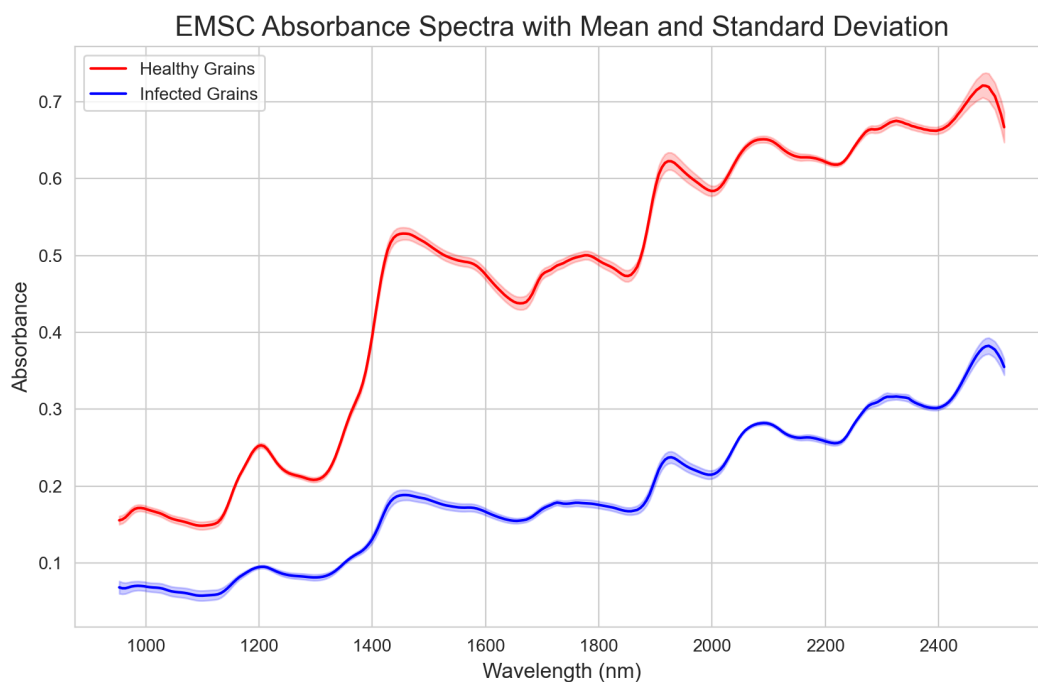
Alternative ML methods were explored to identify wheat grains using hyperspectral data. Four well-known algorithms, including Random Forest (RF), a linear Support Vector Machine (SVM), DecisionTree using the Light Gradient Boosting Machine (LGBM) implementation and a one-vs-all calibrated Stochastic Gradient Descent (SGD) classifier, were trained using all bands from the hyperspectral data.

DecisionTree (LGBM) is a tree-based model that finds the leaves which will reduce the loss the most, and split only that leaf without considering the rest of the leaves in the same level. While RF is an ensemble-based method that constructs multiple decision trees and outputs the class by aggregating the results of each individual tree and selecting the most commonly predicted class as the final output. The linear SVM is a discriminative model that tries to find the best separating hyperplane between two classes. The one-vs-all calibrated SGD classifier is a gradient-based optimization algorithm that updates the model parameters with a small batch of data at a time. It is designed to handle multiclass classification problems by training a separate binary classifier for each class and calibrating the output probabilities. However, for this thesis, the binary version of the one-vs-all calibrated SGD classifier was used, which only considers the healthy class versus the infected class.

The trained models were tested, and their performance was evaluated in terms of precision, recall, and F1-score. This was done to determine whether any of the alternative models could achieve better results than the PLS-DA model.

### 3.6 Aquaphotomics Analysis of Wheat Grains

Samples of wheat grains were prepared and analyzed using the aquaphotomics approach to investigate their molecular composition. Prior to analysis, all samples underwent an EMSC to account for baseline drift and other sources of spectral variation. Following the EMSC procedure, the absorbance spectra of healthy and infected grains were normalized using Eq. (7). The corresponding EMSC absorbance spectra, calculated prior to normalization for visualization in the aquagram, are presented in Fig. 3.12.



**Fig. 3.12:** EMSC absorbance spectra applied before normalizing the values to be plotted in aquagram.

The aquagram was used to compare the absorbance values at different WAMACS positions between healthy and infected grains. In particular, second derivative spectra was used, which allows for the identification and highlighting of specific points of interest within these regions. These points of interest may correspond to transitions or features that may indicate differences in the chemical or structural composition of the grains. Through this approach, specific spectral markers could be identified that offer a more detailed characterization of the differences between healthy and infected grains. The aquagram was generated by plotting the mean normalized absorbance spectra of all samples at each wavelength associated with the 12 distinct WAMACS bands in the first overtone, as listed in Table 2.1.

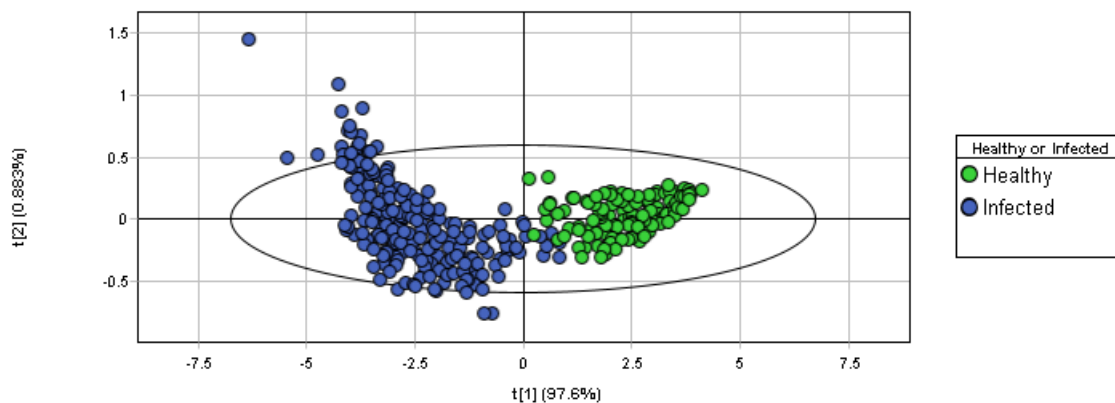
## 4 Results

### 4.1 Classification of the Hyperspectral Data

This subsection focuses on the classification of hyperspectral data using various techniques. It includes an analysis of score variance plot, a comparison between object-based and pixel-based classification methods, and an exploration of the cost-effectiveness of using effective wavelengths for classification.

#### 4.1.1 Score Variance Plot Analysis

Based on the PLS-DA model, the relationship between healthy and infected grains was investigated. To visualize the distribution of the data and detect potential outliers, a score variance plot was generated and presented in Fig. 4.1. The score variance plot is a valuable tool for identifying clusters of data points and extreme values. Consequently, a T2 ellipse was plotted to provide a measure of the confidence interval for the scores. The data points located around the origin represent the region where most of the model's scores are expected to cluster, indicating a high degree of similarity among these samples. This is expected since the origin represents the average of the data used for model training, and the samples that cluster around it have a high level of agreement with the model. Conversely, samples that are located far away from the origin are dissimilar to the model and may indicate potential outliers or samples that do not fit well with the model.



**Fig. 4.1:** Score variance plot to visualize the distribution of data points in the PLS-DA model for healthy and infected grains, and to identify potential outliers.

Fig. 4.1 reveals that all healthy grain variables are well within the T2 ellipse, indicating a higher level of confidence in the PLS model's ability to classify them correctly. However, there are several infected grain data points that fall outside the T2 ellipse, implying that these data points may be considered outliers. Furthermore, the score variance plot depicts that the first latent component explains 97.6% of the total variance in the score plot, indicating that it is a significant contributor to the overall relationship between healthy and infected grains.

### 4.1.2 Object-Based vs. Pixel-Based Classification

Two classification methods were evaluated using a test set, using the last eight images in data set, of 160 grains (80 healthy, 80 infected) and a training set of 640 samples. Both PCA and SNV correction were applied to the data during training. The classification accuracy of the AMS and PM methods was assessed using two confusion matrices (Fig. 4.2 and 4.3). All 288 bands were used for the model, including the first and last bands despite it being common to remove them due to noise. This decision was made based on the loading plot in Fig. 3.7, which showed a significant negative correlation between the principal components of the model. Additionally, breeze has an importance feature for each band during the training process which amplified this statement.

Actual classes	Total	Healthy	Infected
Healthy	80 (50%)	80 (100%)	
Infected	80 (50%)	2 (2.5%)	78 (97.5%)
# Predicted	160 (100%)	82 (51.2%)	78 (48.8%)
Correctly	158 (98.8%)		
Incorrectly	2 (1.25%)		
Precision		97.6%	100%
Recall		100%	97.5%
F-score		98.8%	98.7%

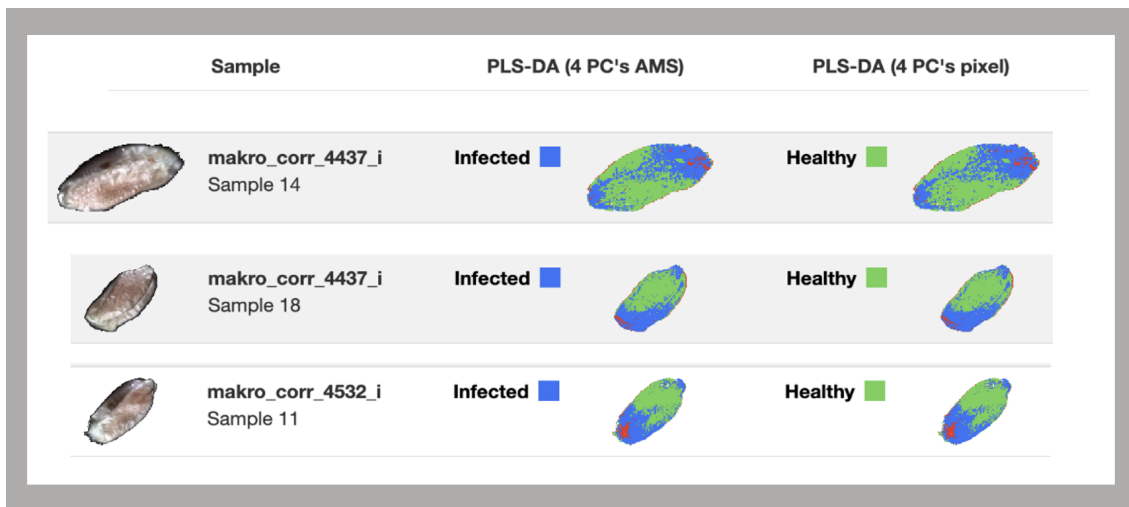
**Fig. 4.2:** Confusion matrix for object-based classification using AMS on test data.

Actual classes	Total	Healthy	Infected
Healthy	80 (50%)	80 (100%)	
Infected	80 (50%)	5 (6.25%)	75 (93.8%)
# Predicted	160 (100%)	85 (53.1%)	75 (46.9%)
Correctly	155 (96.9%)		
Incorrectly	5 (3.12%)		
Precision		94.1%	100%
Recall		100%	93.8%
F-score		97%	96.8%

**Fig. 4.3:** Confusion matrix for pixel-based classification using PM on test data.

The AMS method exhibited a lower FP rate, with two misclassifications where infected grains were detected as healthy. In contrast, the PM method had a higher FP rate, with five misclassifications where infected grains were falsely identified as healthy. These results suggest that the AMS method may be more effective at identifying infected grains with a lower rate of FPs. Appendix A contains the classification report that compares the performance of the PLS-DA model for object-based

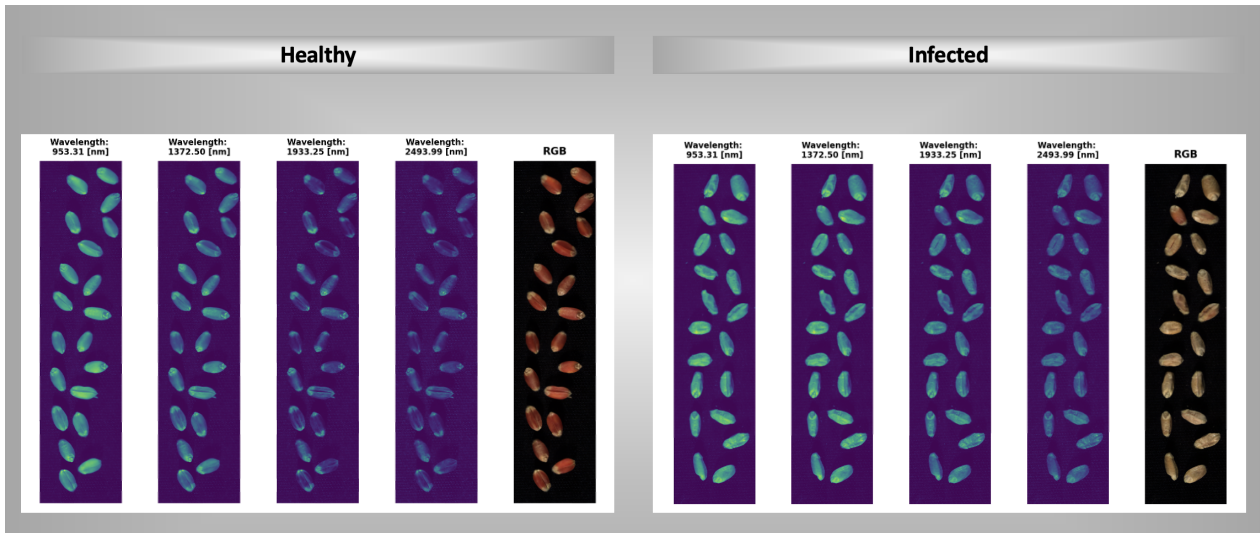
and pixel-based classification, including pixel-wise prediction. The classification report reveals discrepancies between object-based and pixel-based classification. Fig. 4.4 highlights examples of these discrepancies, showing that the PM classifier wrongly classified some grains, while the AMS classifier made correct classifications.



**Fig. 4.4:** Classification examples of infected grains using object-based and pixel-based methods. The PM classifier made incorrect classifications, while the AMS classifier correctly classified the chosen grains.

### 4.1.3 All Wavelengths vs. Effective Wavelengths

The comparison between the object-based classification method using the AMS approach and a model trained on four effective wavelengths was conducted to assess the potential of reducing the number of bands for training and its impact on accuracy. The effectiveness of both methods was evaluated by including all wheat varieties in the analysis. To increase the robustness of the model, four samples were selected from each wheat variety, rather than using only the last eight images as the test set. The images associated with the four effective wavelengths selected based on the loading plot in Fig. 3.7 are depicted in Fig. 4.5.



**Fig. 4.5:** The presented images are composed of various spectral bands, wherein each image corresponds to a specific wavelength within four distinct intervals in addition to the RGB channels. The images on the left represent healthy grains, while those on the right show infected grains.

The results presented in Fig. 4.6 and Fig. 4.7 show that the AMS method achieved near perfect classification accuracy, with only one instance of misclassification, where an infected grain was identified as healthy. In contrast, the model trained on four effective wavelengths had two instances of misclassification, where healthy grains were identified as infected. To further illustrate the differences between the two models, specific grains were selected and are presented in Fig. 4.8, highlighting instances where the classification results differ from the confusion matrices. A classification report for the PLS-DA model predicting all grains in test set, including pixel-wise prediction, is shown in Appendix B for the model trained on effective wavelengths and Appendix C for the model using all wavelengths.

Actual classes	Total	Healthy	Infected
Healthy	80 (50%)	80 (100%)	
Infected	80 (50%)	1 (1.3%)	79 (98.7%)
# Predicted	160 (100%)	81 (50.6%)	79 (49.4%)
Correctly	159 (99.4%)		
Incorrectly	1 (0.6%)		
Precision		98.8%	100%
Recall		100%	98.8%
F-score		99.4%	99.4%

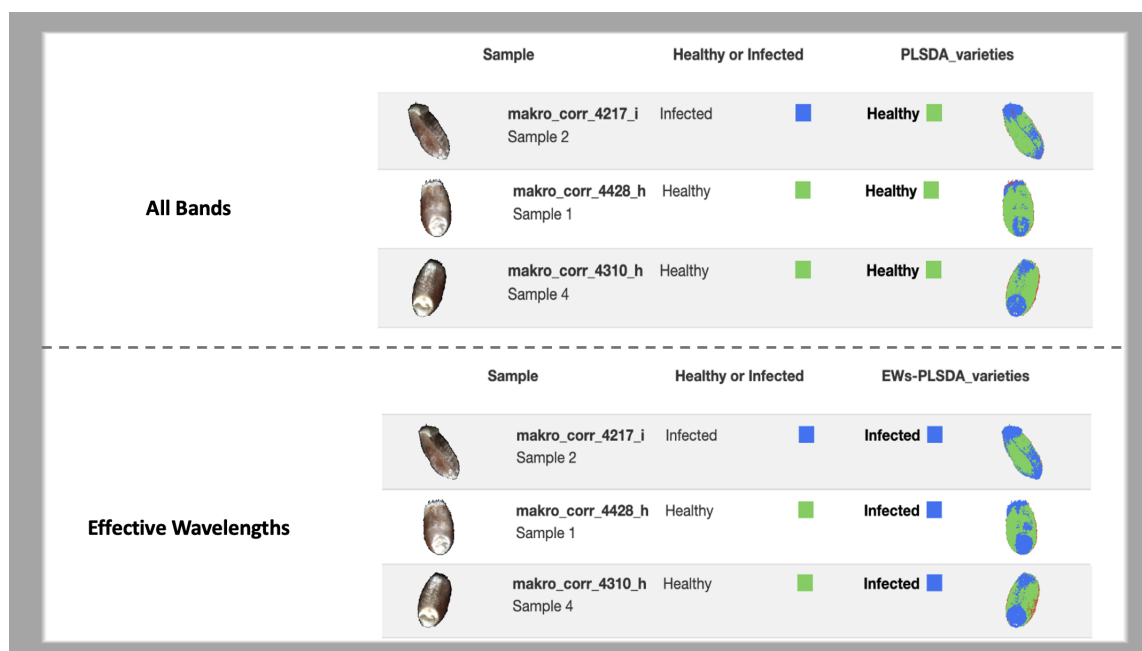
**Fig. 4.6:** Confusion matrix for object-based classification using AMS with all bands on every wheat variety as test data.



Actual classes	Total	Healthy	Infected
Healthy	80 (50%)	78 (97.5%)	2 (2.5%)
Infected	80 (50%)		80 (100%)
# Predicted	160 (100%)	78 (48.8%)	82 (51.2%)
Correctly	158 (98.8%)		
Incorrectly	2 (1.25%)		
Precision		100%	97.6%
Recall		97.5%	100%
F-score		98.7%	98.8%

**Fig. 4.7:** Confusion matrix for object-based classification using AMS with EWs on every wheat variety as test data.

Although the accuracy of the model trained on effective wavelengths was slightly lower than that of the AMS method, the difference was not significant. Therefore, the model trained on four effective wavelengths is a cost-effective and viable alternative to using all bands, which can significantly improve the model's cost-effectiveness and efficiency without compromising accuracy.



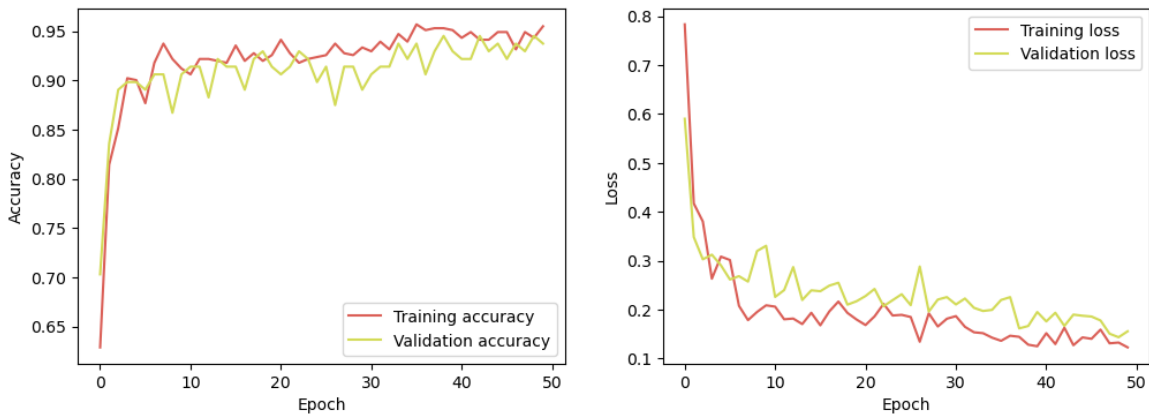
**Fig. 4.8:** Examples of grain classification using EWs and all bands for every wheat variety. The "Healthy or Infected" column represents the ground truth, while the last column displays the prediction.

## 4.2 Deep Learning and Non-Deep Learning Classification

This subsection shows the classification results based on deep learning methods such as CNN for RGB images and multiple ML models for the hyperspectral images.

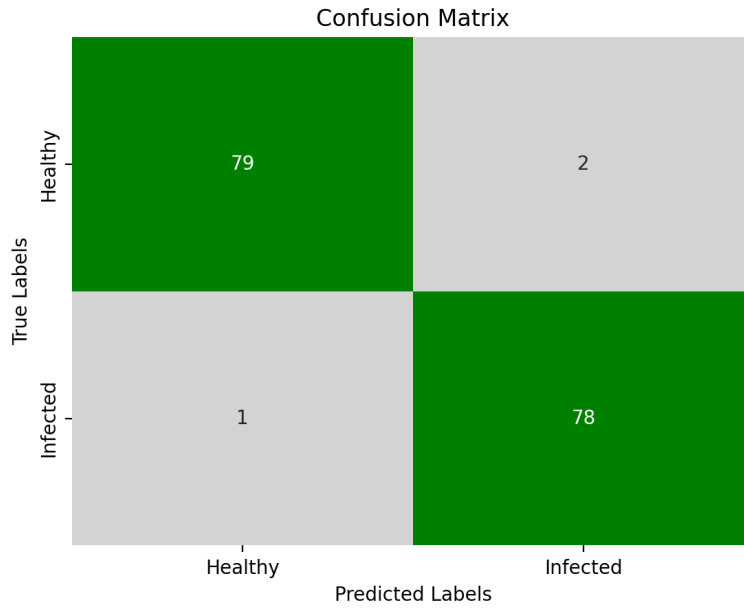
### 4.2.1 CNN-based Analysis with RGB Images

Following the successful implementation of effective wavelengths for identifying wheat grains using hyperspectral images, the possibility of using RGB images was explored as a potential alternative approach. The CNN model was implemented and optimized using the scikit-learn and Keras libraries, respectively. The model architecture consisted of two convolutional layers, each followed by a max-pooling layer and batch normalization, and a fully connected dense layer with dropout included. The number of filters in the convolutional layers, kernel size, number of units in the dense layer, and learning rate were optimized using the GridsearchCV method. The best hyperparameters were found to be: Filters=12, kernel size=3, dense units=18, and learning rate=0.001. The accuracy and loss curves across 50 epochs can be seen in Fig. 4.9.



**Fig. 4.9:** Accuracy and loss curves of the CNN model for training and validation datasets.

Data augmentation improved the generalization performance of the developed CNN model, as evidenced by table 4.1 with a F1 score of 98,1% on the test set. The ImageDataGenerator class from Keras generated new training images, resulting in a larger and more diverse dataset for the model to learn from. The dataset was divided into three subsets, including the train, validation, and test sets, consisting of 512, 128, and 160 images, respectively. Each image had a resolution of 128x128 pixels and consisted of three channels representing the RGB color space, with pixel values normalized to the range [0, 1]. The classification performance of the developed CNN model can be visualized through the confusion matrix in Fig. 4.10. The model achieved an F1-score of 98.1% on the test set, indicating that it was able to accurately classify wheat grains into their respective categories, with only two healthy grains misclassified as infected and one infected grain misclassified as healthy.



**Fig. 4.10:** Confusion matrix predicted by the CNN model for binary classification of wheat grains.

**Table 4.1:** Precision, recall, and F1 score for the CNN model based on the confusion matrix Fig. 4.10.

Condition	Precision	Recall	F1-score
Healthy	0.975	0.988	0.981
Infected	0.987	0.975	0.981

#### 4.2.2 Alternative Machine Learning Algorithms

Although the object-based AMS classifier yielded good results in identifying wheat grains with hyperspectral data, other ML models were investigated to determine whether better results could be achieved. To this end, four commonly used ML models, namely DecisionTree, RF, linear SVM and a SGD calibrated one-vs-all classifier, were evaluated. All models were trained using all bands from the hyperspectral data. The confusion matrices for these models are presented in Fig. 4.11.

DecisionTree (LGBM)				Random Forest			
Actual classes	Total	Healthy	Infected	Actual classes	Total	Healthy	Infected
Healthy	80 (50%)	78 (97.5%)	2 (2.5%)	Healthy	80 (50%)	78 (97.5%)	2 (2.5%)
Infected	80 (50%)	4 (5%)	76 (95%)	Infected	80 (50%)	4 (5%)	76 (95%)
# Predicted	160 (100%)	82 (51.2%)	78 (48.8%)	# Predicted	160 (100%)	82 (51.2%)	78 (48.8%)
Correctly	154 (96.2%)			Correctly	154 (96.2%)		
Incorrectly	6 (3.75%)			Incorrectly	6 (3.75%)		
Precision		95.1%	97.4%	Precision		95.1%	97.4%
Recall		97.5%	95%	Recall		97.5%	95%
F-score		96.3%	96.2%	F-score		96.3%	96.2%

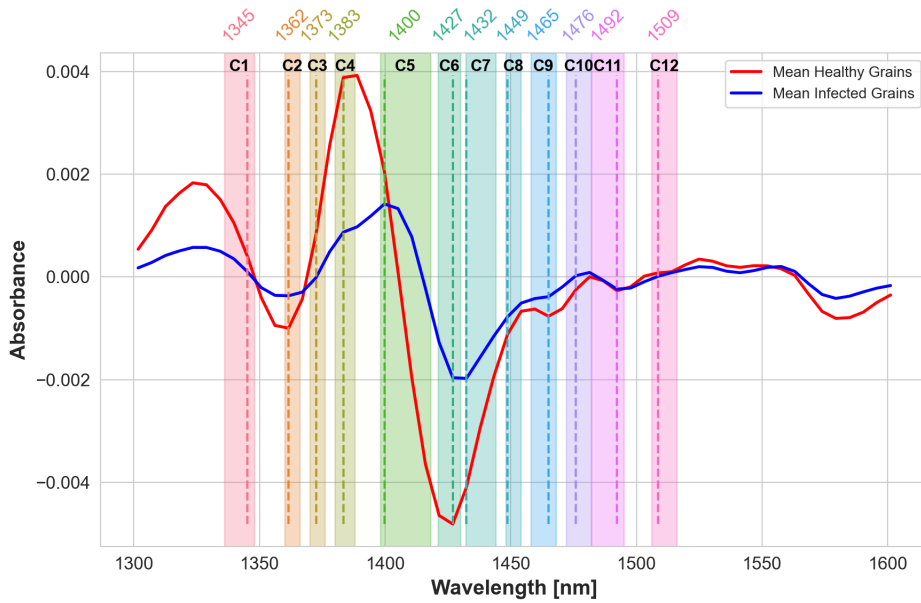
Stochastic Gradient Descent				Support Vector Machine			
Actual classes	Total	Healthy	Infected	Actual classes	Total	Healthy	Infected
Healthy	80 (50%)	80 (100%)		Healthy	80 (50%)	80 (100%)	
Infected	80 (50%)	4 (5%)	76 (95%)	Infected	80 (50%)	7 (8.75%)	73 (91.3%)
# Predicted	160 (100%)	84 (52.5%)	76 (47.5%)	# Predicted	160 (100%)	87 (54.4%)	73 (45.6%)
Correctly	156 (97.5%)			Correctly	153 (95.6%)		
Incorrectly	4 (2.5%)			Incorrectly	7 (4.37%)		
Precision		95%	100%	Precision		92%	100%
Recall		100%	95%	Recall		100%	91.3%
F-score		97.5%	97.4%	F-score		95.8%	95.4%

**Fig. 4.11:** Comparison of DecisionTree (LGBM), RF, one-vs-all SGD, and linear SVM models. Confusion matrices depict the test set results with corresponding precision, recall, and F1-score.

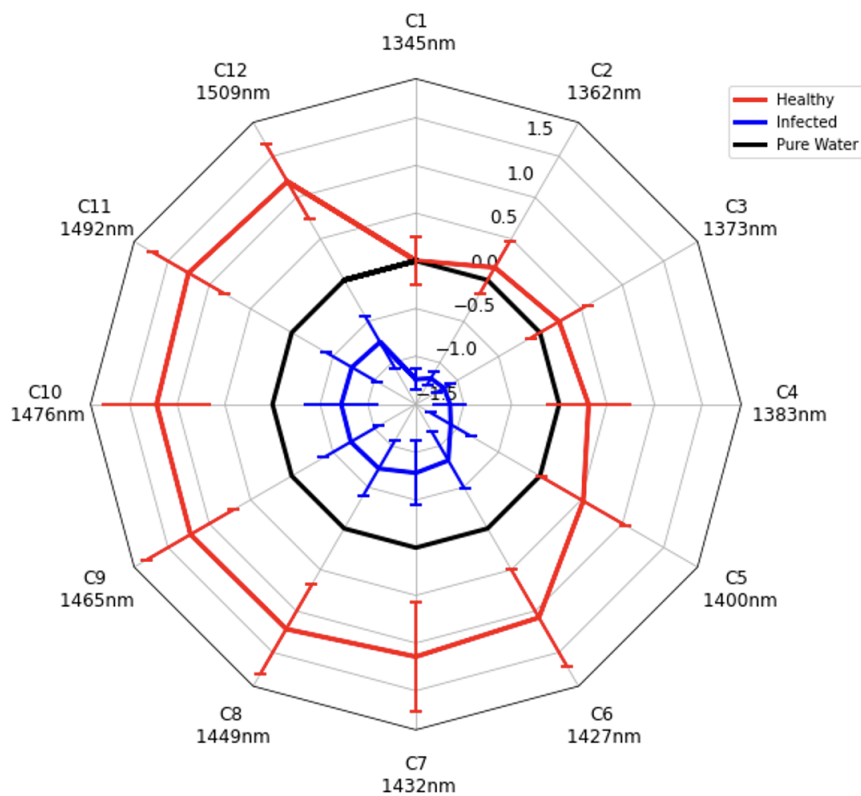
Both the DecisionTree and RF models exhibited two misclassifications of healthy seeds as infected and four misclassifications of infected seeds as healthy, resulting in similar performance. Additionally, SGD and SVM models were tested, but neither model yielded any better results than the AMS approach. The SGD model presented four misclassifications of infected seeds as healthy, while the SVM model had seven such misclassifications.

### 4.3 Moisture and Water Content with Aquaphotomics Analysis

The analysis related to water and moisture content using aquaphotomics revealed distinct differences between healthy and FHB infected wheat grains. Fig. 4.12 shows the mean and corresponding second derivative spectra for healthy and infected grains, indicating that the spectra exhibit similar patterns, except for the 1300-1450 *nm* region, where more variability was observed. The plot also highlights each WAMAC region with a distinct color-coding. The vertical dashed line shows the specific wavelength extracted for plotting in the aquagram Fig. 4.13.



**Fig. 4.12:** Second derivative spectra for healthy and infected grains with their corresponding WAMACS in each region. The dashed vertical line shows the chosen wavelength in each region.



**Fig. 4.13:** Aquagram with WASPs for each WAMAC and pure water.

Descriptive statistics for each of the 12 WAMACS are presented in Table 2.1. The aquagram plot displays the WASPs of each WAMAC ranging from C1 to C12 for both the healthy and infected grains, as well as pure water used as a reference spectrum. The plot highlights that, although

both the infected and healthy grains show a similar pattern, the healthy grains exhibit considerably higher absorbance values, particularly in the C6 to C12 range (hydrogen-bonded molecules and hydration shell), which are responsible for the bending and symmetric stretching modes observed in this range, compared to the infected grains. Furthermore, the absorbance spectra of the healthy grains intersect with pure water at C1 (asymmetrical stretching), while for both the infected and healthy grains, the absorbance values gradually increase from C2 to C5, which differ from C6 to C12, with asymmetrical stretching and no hydrogen-bonded molecules. The error bars in Fig. 4.13 represent the fluctuation range of absorbance values for the grains, corresponding to the standard deviation of the spectra for each grain in the selected WAMAC.

## 5 Discussion

### 5.1 Image Acquisition and Preprocessing

Achieving accurate disease recognition using image-based techniques heavily relies on the quality of the images used [78]. In this regard, the resolution of the images is a crucial factor that plays a significant role in the success of classification methods. However, even after calibration (Eq. 10), the dataset still suffered from significant noise, particularly with regards to the 30 *cm* lens. One possible reason for this is the incorrect camera alignment onto the mobile platform shown in Fig. 3.2 during setup, leading to a mismatch between the focal point and the distance between the camera lens and the grains. Additionally, another potential issue that may have contributed to this noise could be the camera's focus on the petri dish's higher edges, rather than the grains. This, in turn, negatively impacts the clarity and resolution of the images. In an attempt to address this issue, a macro lens was utilized, which resulted in improved image quality and resolution. Although the use of a macro lens was associated with a narrower FOV, the petri dish was switched with a black fabric and the overall improvements in image quality were beneficial in terms of their potential to distinguish between FHB infected and healthy wheat grains.

Before switching to a black fabric as the background for the hyperspectral images, the scanned images suffered from the presence of shadows. In an effort to mitigate this issue, various pre-processing techniques were employed, including SNV transformation (Eq. 12), EMSC (Eq. 14), and first (Eq. 15) and second (Eq. 16) derivatives. Despite the implementation of these methods, as well as other image processing packages such as morphology and color space conversion from RGB to LAB, the images still contained shadows which made it difficult to properly segment out the grains. However, the use of a black fabric as the background effectively eliminated the presence of shadows in the images.

Specular- and scattering reflection of the grains were also a significant challenge. In Fig. 3.3, it can be observed that the reflectance spectra of infected grains were considerably higher than those of the healthy grains. This observation is in line with the findings of previous studies [79, 80], which also reported higher spectral reflectance of fusarium-infected grains compared to healthy ones. These studies utilized non-invasive and non-destructive techniques such as NIR reflectance spectroscopy and HSI to detect fusarium infection in grains. As a result, the infected grains were exposed to higher levels of saturation in the images, making it harder to differentiate between healthy and infected regions.

### 5.2 Classification of Spectral Data

The choice of classification method can significantly impact accuracy in image analysis. Two studies [81, 82] have compared commonly used classification methods, such as RF and SVM, in different landscapes, including pixel-based and object-based classification methods. In this thesis on classifying grains with spectral data, the main model used was PLS-DA. From the results shown

in Fig. 4.2 and 4.3 for the object-based and pixel-based methods, respectively, it can be seen that the AMS classifier yielded better results with fewer infected grains misclassified as healthy. This is consistent with the findings of the previous studies, which have also reported higher accuracy with object-based methods [81].

Although the pixel-based method is quick and easy to use and preserves the spatial details of the image, it may not be as fast as the object-based method. However, the pixel-based method is well-suited for regions with uniform characteristics, but may be more sensitive to noise and image variations, which can lead to incorrect classification. On the other hand, the object-based approach takes into account the spectral characteristics, including shape and texture, is more resilient to noise and variations. However, this method requires more computing power. This difference in methodology may explain why the pixel-based method had difficulty classifying the first grain in Fig. 4.4, as it heavily relies on the dominant color and does not consider the average spectrum. In contrast, the object-based classifier was able to correctly classify the infected grain by considering the average spectrum along with pixel-wise predictions.

### 5.2.1 Loadings

The loadings of the PLS-DA model were examined to identify the wavelengths that contributed the most to the classification of the different classes. As shown in Fig. 3.7, the loadings of the PLS-DA model vary across the spectral region, indicating that certain wavelengths are more important than others for discriminating between the different classes. Fig. 4.5 shows the wavelengths 953 *nm*, 1373 *nm*, 1923 *nm*, and 2493 *nm*, which appears to be the most important for predicting the classes of grains based on the loadings. The negative correlation between the components of these wavelengths indicates that they are likely capturing distinct characteristics of the grains. Using only these four wavelengths could potentially enhance the efficiency and cost-effectiveness of the model. However, it is important to note that the removal of other wavelengths could also lead to loss of information and potentially reduce the accuracy of the model. Furthermore, it is essential to notice that the presence of noise can impact the loadings of the PLS-DA model and their contribution to the classification process. To investigate this, the initial and final ten bands were removed during the PLS-DA model training process. The outcomes demonstrated no improvement compared to using all bands, indicating that the first and last wavelength in the loading plot may have contributed to enhancing the classification results.

### 5.2.2 Effect of Wheat Varieties on Grain Classification

The thesis also investigated the robustness of the PLS-DA model in the context of classification within using every variety of wheat, presented in Table 3.1. Fig. 4.6 and 4.7 shows the results using every wheat variety as test data. The inclusion of every variety in the test set accounted for variations within each type, thereby increasing the reliability of the model. It is worth noting that one of the wheat varieties, namely route 4223, contained white-seeded healthy grains that may have posed difficulties in distinguishing them from infected grains based on color alone. However, this



did not appear to be a problem in the hyperspectral images as the grains appeared darker and were easily distinguishable after preprocessing.

In addition, various commonly used ML algorithms, including DecisionTree, RF, linear SVM, and a one-vs-all calibrated SGD classifier, were applied to the dataset containing all wheat varieties, as illustrated in Fig. 4.11. A recent study on corn seed classification using combined spectral and appearance characteristics reported promising results with SVM outperforming PLS-DA [83]. However, in this particular thesis, the PLS-DA model outperformed the other ML algorithms, suggesting that it is a robust and effective approach for grain classification, particularly when accounting for variations within each wheat variety. The PLS-DA model is particularly well-suited for small sample sizes, due to its ability to handle multicollinearity and to effectively extract relevant information from a limited number of samples [42].

### 5.3 Classification of RGB Images

In the case of RGB images used with a CNN model, the classification accuracy results shown in Fig. 4.10 and Table 4.1 were acceptable. Notably, two healthy grains were misclassified as infected, which is a different pattern from what is observed earlier with the hyperspectral images in Fig. 4.2 and 4.3, where infected grains were misclassified as healthy. This is also consistent with Fig. 4.7 for the PLS-DA model trained on effective wavelengths. The results suggest that the model might have difficulty accurately classifying healthy grains without the additional spectral information, using only three channels for RGB and four bands for effective wavelengths. A potential factor contributing to the misclassification is the white-seeded wheat variety present in the test set. Also, the data for the RGB images only underwent simple normalization techniques, such as scaling the pixel values within the range 0 to 1. While this can help with scaling and centering image data, it may not be sufficient for handling complex variations in lighting like SNV method. Consequently, the model may be more sensitive to brightness variations and wrongly classify healthy grains based on their bright color.

After segmenting the grains, the dataset contained 800 samples. This size is significantly smaller than the datasets used in similar studies. For example, Picon et al. [84] studied wheat diseases using 8178 images, while Lu et al. [85] collected a wheat disease database consisting of 9320 images. The difference in dataset size highlights the advantage of having a larger volume of images for achieving more reliable results when evaluating classification models. Even though data augmentation was used in the training process for the CNN with RGB images, it remains uncertain whether this resulted in an increased robustness of the model, given the limited variations in shape and color observed among the 800 grains.

### 5.4 Moisture and Water Content

The moisture content in seeds is deeply related to their quality and storage life [86]. The WASPs displayed on the aquagram in Fig. 4.13 revealed a noteworthy difference in the absorbance values

between C1 and C3, where C1 (asymmetric stretching) demonstrated lower absorbance values compared to C3 (asymmetric and symmetrical stretching). It also implies that the error bar is the lowest for C1, indicating that the values are more reliable due to the relatively lower variation observed among all grains. These variations in absorbance values could be influenced by several factors, including the concentration or orientation of water molecules, or the presence of other substances that may interact differently with the different modes of vibration [87].

Another important factor determining the WASPs is the temperature, as seen in several other studies [88, 89, 90]. As temperature increases, the thermal energy of the system also increases, which can cause hydrogen bonds between water molecules and other molecules to break [91]. This can have a direct impact on the WASPs. Healthy grains may have stronger hydrogen bonds with the water molecules, which may make them less affected by changes in temperature compared to the infected grains. Consequently, differences in WASPs are observed in the aquagram plot for healthy and infected grains, particularly in the C6 to C12 range, which is responsible for bending and symmetric stretching modes. Thus, the effect of temperature changes can provide valuable information about the physical and chemical properties of grains based on their WASPs.

Although this thesis focused solely on analyzing the first overtone of the absorbance spectra related to water in the NIR range, it would be valuable for future studies to consider examining the second overtone as well. The second overtone, spanning the range of 800-1100 *nm* [92, 93], provides additional information about the water structure and its interactions with solutes. Therefore, investigating in both the first and second overtone regions of water absorbance spectra could provide a more comprehensive understanding of the aqueous system related to each WAMAC.

## 5.5 Limitations and Uncertainties

This thesis assesses the classification performance of both hyperspectral and RGB images for the task of distinguishing between healthy and infected wheat grains. Interestingly, both types of images were found to yield similar classification results. This finding raises the question of whether HSI provides any additional information beyond what is already achieved with RGB images.

There are several potential reasons why both hyperspectral and RGB images produced comparable results in distinguishing between healthy and infected wheat grains. Firstly, the RGB channels can provide valuable information about the color, texture, and shape of objects, which is relevant for this classification task. Secondly, the spectral variability of healthy and infected wheat grains in the hyperspectral data may be limited, meaning that the hyperspectral data may not provide any additional information beyond what is already present in the RGB channels. Additionally, the complexity and noise in hyperspectral data may limit its ability to provide useful information. Lastly, as the training set used in this thesis consisted of a limited number of samples (800 grains), the additional spectral bands in the hyperspectral data may not have added any significant information beyond what is already present in the RGB channels. The restricted dataset size may have resulted in insufficient variation in spectral information, limiting the ability to detect meaningful differences between healthy and infected grains based on spectral signatures.

It is worth noting that the classification task of healthy vs infected wheat grains was accomplished using a supervised learning approach for the PLS-DA model, which has the limitation of being bounded within the training set. This means that the accuracy of the model is heavily dependent on the quality and size of the training dataset. It is unclear how much variation was present in the training dataset and how much of this variation was captured by the model. This factor could potentially impact the generalization performance of the model to new, unseen samples.

Moreover, as previously mentioned, the classification of the wheat grains was achieved through manual annotation, which is a labor-intensive and subjective process that is prone to human error. This limitation may have introduced some degree of bias or inconsistency in the interpretation of the results. Additionally, it is possible that some features or characteristics of the wheat grains may have been overlooked or misinterpreted during the annotation process, which could have affected the accuracy of the classification results. Therefore, it is important to acknowledge the potential impact of the limitations of the manual annotation process on the interpretation of the results.

## 6 Conclusion

The main objective of this thesis was to distinguish between healthy and FHB infected wheat grains using HSI and RGB images. The classification was achieved through the development of a PLS-DA model for hyperspectral images and a CNN model for RGB images. To compare the success of the object-based and pixel-based approaches for classification, the PLS-DA model was developed using both approaches. The results showed that the object-based method was superior in terms of classification accuracy. Other well-known ML algorithms was also evaluated, including RF, linear SVM, SGD calibrated one-vs-all and DecisionTree, and compared their performance with PLS-DA. The PLS-DA model using an object-based approach outperformed the other ML algorithms. Moreover, the water and moisture content in the grains was examined using hyperspectral images in the region of the first overtone of water, represented through an aquagram. The plot highlighted that although both the infected and healthy grains show a similar pattern, the healthy grains exhibit considerably higher absorbance values for all WAMACS.

The performance of the hyperspectral images was compared to that of RGB images, with a test set that did not include every wheat variety. The results showed a minimal difference in classification accuracy between the CNN model and the PLS-DA model, with F1-scores of 98.1% and 98.8%, respectively. To achieve a more cost-effective approach using hyperspectral images, effective wavelengths were tested using four wavelengths corresponding to 953 *nm*, 1373 *nm*, 1923 *nm* and 2493 *nm*. The results showed that the classification accuracy was similar to the full spectral range, with an F1-score of 98.7%. The PLS-DA model using the object-based method yielded better results when tested on all wheat varieties, with an F1-score of 99.4%.

Based on the findings presented in this thesis, it appears that in the specific classification task of distinguishing healthy from FHB infected wheat grains, RGB images may provide a suitable alternative to hyperspectral images. However, it should be emphasized that further experimentation and investigation may be necessary to confirm these results and explore the potential benefits of HSI for other applications. While this study mainly focuses on the chromatic properties of the grains, further studies should also consider to categorize wheat kernels into multiple classes based on factors such as their shapes, textures or even different stages of FHB-infection severity, which could provide more detailed insights into the characteristics of different varieties of grains. These additional features could benefit the CNN model if used with hyperspectral images as more data is available to differentiate between healthy and FHB infected wheat grains. Additionally, deeper investigations into how water absorption affects spectral measurements and moisture content in grains require further research. Supplementary to this, future studies could also consider exploring the development of a python graphical user interface for a mobile app that allows users to interact with a saved CNN model for image classification. Such an interface could make the implementation of image classification using deep learning models more efficient and user-friendly.

## References

- [1] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, and et al., “A review on deep learning techniques applied to semantic segmentation,” *arXiv preprint arXiv:1704.06857*, 2017.
- [2] H. Bansal, B. Balusamy, T. Poongodi, and F.K. KP, *Machine Learning and Analytics in Healthcare Systems: Principles and Applications*, Green Engineering and Technology. CRC Press, 2021.
- [3] Andreas Vidman and Oskar Jonsson, “Breeze - hyperpectral imaging software,” *Prediktera*, 2015, downloaded 6th march, 2023.
- [4] John C. Avise and Francisco J. Ayala, *In the Light of Evolution: Volume I: Adaptation and Complex Design*, The National Academies Press, Washington, DC, 2007.
- [5] Diego Soto-Gómez and Paula Pérez-Rodríguez, “Sustainable agriculture through perennial grains: Wheat, rice, maize, and other species. a review,” *Agriculture, Ecosystems & Environment*, vol. 325, pp. 107747, 2022.
- [6] FAO, *Food Outlook - Biannual Report on Global Food Markets*, Food and Agriculture Organization of the United Nations, Rome, Italy, 2022, Accessed on March 21, 2023.
- [7] M.P. Reynolds and H.J. Braun, *Wheat Improvement: Food Security in a Changing Climate*, Springer International Publishing, 2022.
- [8] Elżbieta Mielniczuk and Barbara Skwaryło-Bednarz, “Fusarium head blight, mycotoxins and strategies for their reduction,” *Agronomy*, vol. 10, no. 4, 2020.
- [9] Ruicheng Qiu, Ce Yang, Ali Moghimi, Man Zhang, Brian Steffenson, and Cory Hirsch, “Detection of fusarium head blight in wheat using a deep neural network and color imaging,” *Remote Sensing*, vol. 11, pp. 2658, 11 2019.
- [10] Lv Yipeng, Lv Wenbing, Han Kaixuan, Tao Wentao, Zheng Ling, Weng Shizhuang, and Huang Linsheng, “Determination of wheat kernels damaged by fusarium head blight using monochromatic images of effective wavelengths from hyperspectral imaging coupled with an architecture self-search deep network,” *Food Control*, vol. 135, pp. 108819, 2022.
- [11] Guihua Bai and Gregory Shaner, “Management and resistance in wheat and barley to fusarium head blight,” *Annual Review of Phytopathology*, vol. 42, no. 1, pp. 135–161, 2004, PMID: 15283663.
- [12] Anis Ben Amar, Souheib Oueslati, Abdelwahed Ghorbel, and Ahmed Mliki, “Prediction and early detection of mycotoxigenic fusarium culmorum in wheat by direct pcr-based procedure,” *Food Control*, vol. 23, no. 2, pp. 506–510, 2012.

- [13] C. Maragos, M. Busman, and Y. Sugita-Konishi, “Production and characterization of a monoclonal antibody that cross-reacts with the mycotoxins nivalenol and 4-deoxynivalenol,” *Food Additives & Contaminants*, vol. 23, no. 8, pp. 816–825, 2006, PMID: 16807207.
- [14] Ali Atoui, André El Khoury, Mireille Kallassy, and Ahmed Lebrihi, “Quantification of fusarium graminearum and fusarium culmorum by real-time pcr system and zearalenone assessment in maize,” *International Journal of Food Microbiology*, vol. 154, no. 1, pp. 59–65, 2012.
- [15] Senay Simsek, Kimberly Burgess, Kristin L. Whitney, Yan Gu, and Steven Y. Qian, “Analysis of deoxynivalenol and deoxynivalenol-3-glucoside in wheat,” *Food Control*, vol. 26, no. 2, pp. 287–292, 2012.
- [16] Antonia J. Powell and Vladimir Vujanovic, “Evolution of fusarium head blight management in wheat: Scientific perspectives on biological control agents and crop genotypes protocooperation,” *Applied Sciences*, vol. 11, no. 19, 2021.
- [17] Jordan R Ubbens and Ian Stavness, “Deep plant phenomics: A deep learning platform for complex plant phenotyping tasks,” *Frontiers in plant science*, vol. 8, pp. 1190, 2017.
- [18] Annika Jäkel, “Automation of the labeling of images of sugar beet cultivation with hyperspectral imaging,” *Norwegian University of Life Sciences*, 2020.
- [19] Ruicheng Qiu, Ce Yang, Ali Moghimi, Man Zhang, Brian J. Steffenson, and Cory D. Hirsch, “Detection of fusarium head blight in wheat using a deep neural network and color imaging,” *Remote Sensing*, vol. 11, no. 22, 2019.
- [20] Peng Zhao, Shanshan Gu, Chongjing Han, Yanli Lu, Changjie Ma, Jingjing Tian, Jinhua Bi, Zhiyong Deng, Qi Wang, and Qiang Xu, “Targeted and untargeted metabolomics profiling of wheat reveals amino acids increase resistance to fusarium head blight,” *Frontiers in Plant Science*, vol. 12, pp. 762605, 2021.
- [21] Cristina Malegori, Jelena Muncan, Eleonora Mustorgi, Roumiana Tsenkova, and Paolo Oliveri, “Analysing the water spectral pattern by near-infrared spectroscopy and chemometrics as a dynamic multidimensional biomarker in preservation: rice germ storage monitoring,” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 265, pp. 120396, 2022.
- [22] Stefka Atanassova, “Near infrared spectroscopy and aquaphotomics for monitoring changes during yellow cheese ripening,” *Agricultural Science and Technology*, vol. 7, pp. 269–272, 06 2015.
- [23] C. Tominski and H. Schumann, *Interactive Visual Data Analysis*, AK Peters Visualization Series. CRC Press, 2020.
- [24] S. De, S. Bhattacharyya, S. Chakraborty, and P. Dutta, *Hybrid Soft Computing for Multi-level Image and Data Segmentation*, Computational Intelligence Methods and Applications. Springer International Publishing, 2016.

- [25] Chanki Pandey, Yogesh Sahu, Prabira Sethy, and Santi Behera, *Hyperspectral Imagery Applications for Precision Agriculture*, pp. 1–38, 02 2022.
- [26] Chunying Wang, Baohua Liu, Lipeng Liu, Yanjun Zhu, Jialin Hou, Ping Liu, and Xiang Li, “A review of deep learning used in the hyperspectral image analysis for agriculture,” *Artificial Intelligence Review*, vol. 54, no. 7, pp. 5205–5253, 2021.
- [27] F. Vasefi, N. MacKinnon, and D.L. Farkas, “Chapter 16 - hyperspectral and multispectral imaging in dermatology,” in *Imaging in Dermatology*, Michael R. Hamblin, Pinar Avci, and Gaurav K. Gupta, Eds., pp. 187–201. Academic Press, Boston, 2016.
- [28] Yang Tang, Shuang Song, Shengxi Gui, Weilun Chao, Chinmin Cheng, and Rongjun Qin, “Active and low-cost hyperspectral imaging for the spectral analysis of a low-light environment,” *Sensors*, vol. 23, no. 3, 2023.
- [29] Jonghee Yoon, “Hyperspectral imaging for clinical applications,” *BioChip Journal*, vol. 16, 01 2022.
- [30] J.M. Amigo, *Hyperspectral Imaging*, ISSN. Elsevier Science, 2019.
- [31] V. F Weisskopf, “How light interacts with matter,” *Scientific American*, vol. 219, no. 5, pp. 60–71, 2010, downloaded 19th march, 2023.
- [32] M. Milosevic, *Internal Reflection and ATR Spectroscopy*, Chemical Analysis: A Series of Monographs on Analytical Chemistry and Its Applications. Wiley, 2012.
- [33] G. Franceschetti and D. Riccio, *Scattering, Natural Surfaces, and Fractals*, Elsevier Science, 2006.
- [34] José Manuel Amigo, Hamid Babamoradi, and Saioa Elcoroaristizabal, “Hyperspectral image analysis. a tutorial,” *Analytica Chimica Acta*, vol. 896, pp. 34–51, 2015.
- [35] S. Raschka and V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*, Packt Publishing, 2019.
- [36] G. Paliouras, V. Karkaletsis, and C.D. Spyropoulos, *Machine Learning and Its Applications: Advanced Lectures*, Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2001.
- [37] Dante Pirouz, “An overview of partial least squares,” *SSRN Electronic Journal*, pp. 3–10, 10 2006.
- [38] Fabian Dablander, “Partial least squares,” *Towards Data Science*, 2021, Downloaded 21. march, 2023.
- [39] Talha Naveed, “Explore the effect of data augmentation of spectroscopic data for deep learning models,” *Master’s Thesis, Norwegian University of Life Sciences*, 2022.
- [40] S. Mishra, U. Sarkar, S. Taraphder, S. Datta, D. Swain, R. Saikhom, S. Panda, and M. Laishram, “Principal component analysis,” *International Journal of Livestock Research*, vol. 1, 2017.

- [41] J. Brownlee, *Master Machine Learning Algorithms: Discover How They Work and Implement Them From Scratch*, Machine Learning Mastery, 2016.
- [42] V.E. Vinzi, W.W. Chin, J. Henseler, and H. Wang, *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer Handbooks of Computational Statistics. Springer Berlin Heidelberg, 2010.
- [43] Z.R. Yang and Z. Yang, “6.01 - artificial neural networks,” in *Comprehensive Biomedical Physics*, Anders Brahme, Ed., pp. 1–17. Elsevier, Oxford, 2014.
- [44] Masoud Monjezi and Hesam Dehghani, “Evaluation of effect of blasting pattern parameters on back break using neural networks,” *International Journal of Rock Mechanics and Mining Sciences*, vol. 45, pp. 1446–1453, 12 2008.
- [45] Colin Reeves and Stewart Taylor, ,” in *Selection of Training Data for Neural Networks by a Genetic Algorithm.*, 09 1998, vol. 1498, pp. 633–642.
- [46] Xue Ying, “An overview of overfitting and its solutions,” *Journal of Physics: Conference Series*, vol. 1168, no. 2, pp. 022022, feb 2019.
- [47] A. Belhi, A. Bouras, A.K. Al-Ali, and A.H. Sadka, *Data Analytics for Cultural Heritage: Current Trends and Concepts*, Springer International Publishing, 2021.
- [48] Junyi Chai, Hao Zeng, Anming Li, and Eric W.T. Ngai, “Deep learning in computer vision: A critical review of emerging techniques and application scenarios,” *Machine Learning with Applications*, vol. 6, pp. 100134, 2021.
- [49] Naiyan Wang, Siyi Li, Abhinav Gupta, and Dit-Yan Yeung, “Transferring rich feature hierarchies for robust visual tracking,” 2015.
- [50] Linwei Ye, Zhi Liu, and Yang Wang, “Dual convolutional lstm network for referring image segmentation,” *IEEE Transactions on Multimedia*, vol. 22, no. 12, pp. 3224–3235, 2020.
- [51] Tianyang Xu, Zhen-Hua Feng, Xiao-Jun Wu, and Josef Kittler, “Joint group feature selection and discriminative filter learning for robust visual object tracking,” in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 7949–7959.
- [52] F. Millstein, *Convolutional Neural Networks In Python: Beginner’s Guide To Convolutional Neural Networks In Python*, Frank Millstein, 2020.
- [53] S. Li, C. Liu, and Y. Wang, *Pattern Recognition: 6th Chinese Conference, CCPR 2014, Changsha, China, November 17-19, 2014. Proceedings, Part II*, Communications in Computer and Information Science. Springer Berlin Heidelberg, 2014.
- [54] I. Zafar, G. Tzanidou, R. Burton, N. Patel, and L. Araujo, *Hands-On Convolutional Neural Networks with TensorFlow: Solve computer vision problems with modeling in TensorFlow and Python*, Packt Publishing, 2018.



- [55] A.E.P. Villa, P. Masulli, and A.J.P. Rivero, *Artificial Neural Networks and Machine Learning – ICANN 2016: 25th International Conference on Artificial Neural Networks, Barcelona, Spain, September 6-9, 2016, Proceedings, Part II*, Lecture Notes in Computer Science. Springer International Publishing, 2016.
- [56] Sergey Ioffe and Christian Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” 2015.
- [57] M. Suri, *Advances in Neuromorphic Hardware Exploiting Emerging Nanoscale Devices*, Cognitive Systems Monographs. Springer India, 2017.
- [58] Xavier Glorot, Antoine Bordes, and Y. Bengio, ,” in *Deep Sparse Rectifier Neural Networks*, 01 2010, vol. 15.
- [59] Q. Liang, W. Wang, J. Mu, X. Liu, and Z. Na, *Artificial Intelligence in China: Proceedings of the 4th International Conference on Artificial Intelligence in China*, Lecture Notes in Electrical Engineering. Springer Nature Singapore, 2023.
- [60] Cheng-Fu Huang, Ding-Hsiang Huang, Yi-Kuei Lin, and Yi-Fan Chen, “Network reliability evaluation of manufacturing systems by using a deep learning approach,” *Annals of Operations Research*, 09 2022.
- [61] A.J. Kulkarni and S.C. Satapathy, *Optimization in Machine Learning and Applications*, Algorithms for Intelligent Systems. Springer Singapore, 2019.
- [62] Dennis Banga and Peter Wagacha, “Abnormality detection in musculoskeletal radiographs with convolutional neural networks(ensembles) and performance optimization,” 08 2019.
- [63] J. Brownlee, *Imbalanced Classification with Python: Better Metrics, Balance Skewed Classes, Cost-Sensitive Learning*, Machine Learning Mastery, 2020.
- [64] Kate Sendin, *Reduced wavelength spectral imaging for grading defect and asymptomatic Fusarium detection in white maize*, Doctoral dissertation, Stellenbosch University, Stellenbosch, South Africa, March 2020, Supervisor: Dr Paul J. Williams, Co-supervisor: Prof. Marena Manley.
- [65] Radmila Tomovska, Jelena Munćan, Ivana Mileusnić, Jovana Šakota Rosić, Aleksandra Vasić-Milovanović, and Lidija Matija, “Water properties of soft contact lenses: A comparative near-infrared study of two hydrogel materials,” *International Journal of Polymer Science*, vol. 2016, pp. 3737916, 2016.
- [66] Roman Maniewski, Renata Meżyk-Kopeć, and Andrzej Wójtowicz, “Aquaphotomics: Water in the biological and aqueous world scrutinised with invisible light,” *Spectroscopy Europe/World*, vol. 22, no. 6, pp. 14–17, 12 2010.
- [67] Noriko Goto, Gyorgy Bazar, Zoltan Kovacs, Makoto Kunisada, Hiroyuki Morita, Seiichiro Kizaki, Hiroshi Sugiyama, Roumiana Tsenkova, and Chikako Nishigori, “Detection of uv-induced cyclobutane pyrimidine dimers by near-infrared spectroscopy and aquaphotomics,” *Scientific Reports*, vol. 5, no. 1, pp. 11808, 2015.

- [68] Roumiana Tsenkova, Jelena Munćan, Bernhard Pollner, and Zoltan Kovacs, “Essentials of aquaphotomics and its chemometrics approaches,” *Frontiers in Chemistry*, vol. 6, 2018.
- [69] Floyd W Nutter, Mark L Gleason, John H Jenco, and Neal C Christians, “Assessing the accuracy, intra-rater repeatability, and inter-rater reliability of disease assessment systems,” *Phytopathology*, vol. 83, pp. 806e812, 1993.
- [70] K.A. Bakeev, *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*, Wiley, 2010.
- [71] Daniel Pelliccia, “Two scatter correction techniques for nir spectroscopy in python,” *NIRPY Research*, July 2018, downloaded 19th march, 2023.
- [72] Nils Kristian Afseth and Achim Kohler, “Extended multiplicative signal correction in vibrational spectroscopy, a tutorial,” *Chemometrics and Intelligent Laboratory Systems*, vol. 117, pp. 92–99, 2012.
- [73] X. Chu, Y. Huang, Y.H. Yun, and X. Bian, *Chemometric Methods in Analytical Spectroscopy Technology*, Springer Nature Singapore, 2022.
- [74] Nils Kristian Afseth and Achim Kohler, “Extended multiplicative signal correction in vibrational spectroscopy, a tutorial,” *Chemometrics and Intelligent Laboratory Systems*, vol. 117, pp. 92–99, 2012, Special Issue Section: Selected Papers from the 1st African-European Conference on Chemometrics, Rabat, Morocco, September 2010 Special Issue Section: Pre-processing methods Special Issue Section: Spectroscopic imaging.
- [75] E. Li-Chan, J.M. Chalmers, and P.R. Griffiths, *Applications of Vibrational Spectroscopy in Food Science, 2 Volume Set*, Wiley, 2010.
- [76] H.G. Jones and R.A. Vaughan, *Remote Sensing of Vegetation: Principles, Techniques, and Applications*, OUP Oxford, 2010.
- [77] M.J. Adams, N.W. Barnett, and Royal Society of Chemistry (Great Britain), *Chemometrics in Analytical Spectroscopy*, RSC analytical spectroscopy monographs. Royal Society of Chemistry, 2004.
- [78] Vajira Thambawita, Inga Strümke, Steven A Hicks, Pål Halvorsen, Sravanthi Parasa, and Michael A Riegler, “Impact of image resolution on deep learning performance in endoscopy image classification: An experimental study using a large dataset of endoscopic images,” *Diagnostics (Basel)*, vol. 11, no. 12, pp. 2183, Nov. 2021.
- [79] Jongguk Lim, Giyoung Kim, Changyeun Mo, Kyoungmin Oh, Hyeonchae Yoo, Hyeonheui Ham, and Moon S. Kim, “Classification of fusarium-infected korean hulled barley using near-infrared reflectance spectroscopy and partial least squares discriminant analysis,” *Sensors*, vol. 17, no. 10, 2017.



















































- [80] Jongguk Lim, Giyoung Kim, Changyeun Mo, Kyoungmin Oh, Geonseob Kim, Hyeonheui Ham, Seongmin Kim, and Moon S. Kim, “Application of near infrared reflectance spectroscopy for rapid and non-destructive discrimination of hulled barley, naked barley, and wheat contaminated with fusarium,” *Sensors*, vol. 18, no. 1, 2018.
- [81] Dennis Duro, Steven Franklin, and Monique Dubé, “A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using spot-5 hrg imagery,” *Remote Sensing of Environment*, vol. 118, pp. 259–272, 03 2012.
- [82] N. Wu, L.G.T. Crusiol, G. Liu, D. Wuyun, and G. Han, “Comparing machine learning algorithms for pixel/object-based classifications of semi-arid grassland in northern china using multisource medium resolution imageries,” *Remote Sens.*, vol. 15, no. 3, pp. 750, 2023.
- [83] Xiaoling Yang, Hanmei Hong, Zhaohong You, and Fang Cheng, “Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification,” *Sensors (Basel, Switzerland)*, vol. 15, pp. 15578–15594, 07 2015.
- [84] Artzai Picon, Aitor Alvarez-Gila, Maximilian Seitz, Amaia Ortiz-Barredo, Jone Echazarra, and Alexander Johannes, “Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild,” *Computers and Electronics in Agriculture*, vol. 161, pp. 280–290, 2019, BigData and DSS in Agriculture.
- [85] Jiang Lu, Jie Hu, Guannan Zhao, Fenghua Mei, and Changshui Zhang, “An in-field automatic wheat disease diagnosis system,” *Computers and Electronics in Agriculture*, vol. 142, pp. 369–379, 2017.
- [86] Muhammad Aqeel Bakhtavar, Imran Afzal, and Shahzad Maqsood Ahmad Basra, “Moisture adsorption isotherms and quality of seeds stored in conventional packaging materials and hermetic super bag,” *PLoS One*, vol. 14, no. 2, pp. e0207569, Feb 2019.
- [87] Muna E. Raypah, Ahmad Fairuz Omar, Jelena Muncan, Musfirah Zulkurnain, and Abdul Rahman Abdul Najib, “Identification of stingless bee honey adulteration using visible-near infrared spectroscopy combined with aquaphotomics,” *Molecules*, vol. 27, no. 7, 2022.
- [88] A.C. Peinado, F. van den Berg, M. Blanco, and R. Bro, “Temperature-induced variation for nir tensor-based calibration,” *Chemometrics and Intelligent Laboratory Systems*, vol. 83, no. 1, pp. 75–82, 2006.
- [89] Hisashi Maeda, Yukihiro Ozaki, Munehiro Tanaka, Nobuyuki Hayashi, and Takayuki Kojima, “Near infrared spectroscopy and chemometrics studies of temperature-dependent spectral variations of water: Relationship between spectral changes and hydrogen bonds,” *Journal of Near Infrared Spectroscopy*, vol. 3, no. 4, pp. 191–201, 1995.
- [90] Jun Kang, Wensheng Cai, and Xueguang Shao, “Quantitative determination by temperature dependent near-infrared spectra: A further study,” *Talanta*, vol. 85, no. 1, pp. 420–424, 2011.

- [91] Shiyang Wang, Mian Wang, Li Han, Yan Sun, Wensheng Cai, and Xueguang Shao, “Insight into the stability of protein in confined environment through analyzing the structure of water by temperature-dependent near-infrared spectroscopy,” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 267, pp. 120581, 2022.
- [92] Damenraj Rajkumar, Rainer Künnemeyer, Harpreet Kaur, Jevon Longdell, and Andrew McGlone, “Interactions of linearly polarized and unpolarized light on kiwifruit using aquaphotomics,” *Molecules*, vol. 27, no. 2, 2022.
- [93] Harpreet Kaur, *Investigating aquaphotomics for fruit quality assessment*, Doctor of philosophy (phd), The University of Waikato, Hamilton, New Zealand, 2020.













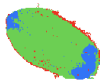

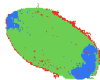



























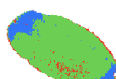







# A Appendix 1 - Limited Wheat Varieties in Test Set



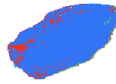

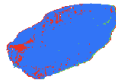

















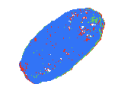

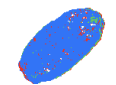


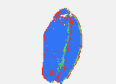

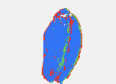




















## Wheat Grains - Daniel.S.L - Test Set

Created 2023-03-16

Time	Sample	PLS-DA (4 PC's AMS)	PLS-DA (4 PC's pixel)
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 1	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 2	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 3	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 4	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 5	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 6	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 7	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 8	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 9	Healthy  	Healthy  
2023-03-14 11:21:24	 <b>makro_corr_4437_h</b> Sample 10	Healthy  	Healthy  



















































**Fig. A.1:** Object-based vs. Pixel-based classification with test set containing only a few wheat varieties. Page 1/16



















































2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 11	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 12	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 13	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 14	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 15	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 16	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 17	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 18	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 19	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 20	<b>Healthy</b> 		<b>Healthy</b> 	



















































2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 1	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 2	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 3	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 4	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 5	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 6	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 7	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 8	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 9	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 10	<b>Infected</b> 		<b>Infected</b> 	



















































2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 11	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 12	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 13	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 14	<b>Infected</b>		<b>Healthy</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 15	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 16	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 17	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 18	<b>Infected</b>		<b>Healthy</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 19	<b>Infected</b>		<b>Infected</b>	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 20	<b>Infected</b>		<b>Infected</b>	






































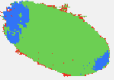

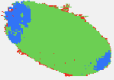






























































2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 1	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 2	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 3	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 4	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 5	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 6	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 7	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 8	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 9	Healthy 		Healthy 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 10	Healthy 		Healthy 	



















































2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 11	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 12	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 13	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 14	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 15	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 16	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 17	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 18	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 19	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 20	<b>Healthy</b> 		<b>Healthy</b> 	



















































2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 1	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 2	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 3	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 4	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 5	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 6	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 7	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 8	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 9	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 10	<b>Infected</b> 		<b>Infected</b> 	

2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 11	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 12	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 13	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 14	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 15	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 16	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 17	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 18	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 19	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 20	<b>Infected</b> 		<b>Infected</b> 	



















































2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 1	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 2	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 3	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 4	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 5	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 6	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 7	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 8	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 9	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 10	<b>Healthy</b> 		<b>Healthy</b> 	

































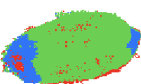

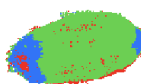















2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 11	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 12	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 13	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 14	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 15	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 16	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 17	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 18	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 19	Healthy 		Healthy 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 20	Healthy 		Healthy 	



















































2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 1	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 2	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 3	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 4	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 5	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 6	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 7	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 8	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 9	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 10	<b>Infected</b> 		<b>Infected</b> 	










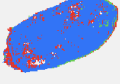



























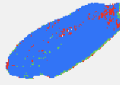

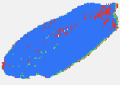


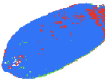

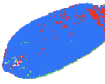





2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 11	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 12	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 13	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 14	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 15	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 16	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 17	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 18	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 19	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 20	<b>Infected</b> 		<b>Infected</b> 	



2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 1	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 2	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 3	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 4	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 5	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 6	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 7	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 8	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 9	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 10	<b>Healthy</b> 		<b>Healthy</b> 	

2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 11	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 12	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 13	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 14	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 15	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 16	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 17	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 18	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 19	Healthy 		Healthy 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 20	Healthy 		Healthy 	



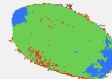










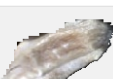









2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 1	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 2	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 3	<b>Healthy</b> 		<b>Healthy</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 4	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 5	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 6	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 7	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 8	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 9	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 10	<b>Infected</b> 		<b>Infected</b> 	

2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 11	<b>Infected</b> 		<b>Healthy</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 12	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 13	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 14	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 15	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 16	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 17	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 18	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 19	<b>Infected</b> 		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 20	<b>Infected</b> 		<b>Infected</b> 	

















































## B Appendix 2 - Every Variety - Effective Wavelengths












































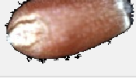



### Varieties Wheat Grains - EWs

















Created 2023-04-03 10:34

Time	Sample	Healthy or Infected	EWs-PLSDA_varieties
2023-03-14 11:01:22	 <b>makro_corr_4205_h</b> Sample 1	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:22	 <b>makro_corr_4205_h</b> Sample 2	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:22	 <b>makro_corr_4205_h</b> Sample 3	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:22	 <b>makro_corr_4205_h</b> Sample 4	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:49	 <b>makro_corr_4205_i</b> Sample 1	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:01:49	 <b>makro_corr_4205_i</b> Sample 2	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:01:49	 <b>makro_corr_4205_i</b> Sample 3	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:01:49	 <b>makro_corr_4205_i</b> Sample 4	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:02:18	 <b>makro_corr_4208_h</b> Sample 1	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:02:18	 <b>makro_corr_4208_h</b> Sample 2	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:02:18	 <b>makro_corr_4208_h</b> Sample 3	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:02:18	 <b>makro_corr_4208_h</b> Sample 4	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 









**Fig. B.1:** Object-based classification using EWs with test set containing every wheat variety. The third column "Healthy or Infected" is the true label, while the predicted label is under the fourth column "EWs-PLSDA\_varieties". Page 1/13

2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 4	Infected		<b>Infected</b> 	










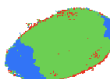



































2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 4	Healthy		<b>Healthy</b> 	




















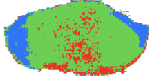




























2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 4	Infected		<b>Infected</b> 	
































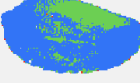


































































2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 4	Healthy		<b>Healthy</b> 	













































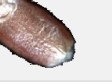



2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 4	Infected		<b>Infected</b> 	

















































2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 4	Healthy		<b>Infected</b> 	

2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 4	Infected		<b>Infected</b> 	




































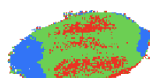




















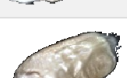







2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 4	Healthy		<b>Healthy</b> 	

2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 4	Infected		<b>Infected</b> 	

2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 1	Healthy		<b>Infected</b> 	
2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 4	Healthy		<b>Healthy</b> 	

2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 4	Infected		<b>Infected</b> 	




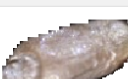










2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 4	Infected		<b>Infected</b> 	

















































## C Appendix 3 - Every Variety - All Bands




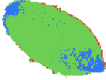










































### Varieties Wheat Grains - All Bands






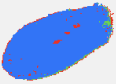








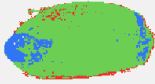

















Created 2023-03-29 15:20











































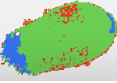
Time	Sample	Healthy or Infected	PLSDA_varieties
2023-03-14 11:01:22	 makro_corr_4205_h Sample 1	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:22	 makro_corr_4205_h Sample 2	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:22	 makro_corr_4205_h Sample 3	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:22	 makro_corr_4205_h Sample 4	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:01:49	 makro_corr_4205_i Sample 1	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:01:49	 makro_corr_4205_i Sample 2	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:01:49	 makro_corr_4205_i Sample 3	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:01:49	 makro_corr_4205_i Sample 4	Infected <span style="color: blue;">■</span>	<b>Infected</b> <span style="color: blue;">■</span> 
2023-03-14 11:02:18	 makro_corr_4208_h Sample 1	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:02:18	 makro_corr_4208_h Sample 2	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:02:18	 makro_corr_4208_h Sample 3	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 
2023-03-14 11:02:18	 makro_corr_4208_h Sample 4	Healthy <span style="color: green;">■</span>	<b>Healthy</b> <span style="color: green;">■</span> 








**Fig. C.1:** Object-based classification using all bands with test set containing every wheat variety. The third column "Healthy or Infected" is the true label, while the predicted label is under the fourth column "PLSDA\_varieties". Page 1/13












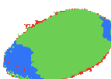




































2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:02:46		<b>makro_corr_4208_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:12		<b>makro_corr_4217_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 2	Infected		<b>Healthy</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:03:39		<b>makro_corr_4217_i</b> Sample 4	Infected		<b>Infected</b> 	

2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:05		<b>makro_corr_4222_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:04:31		<b>makro_corr_4222_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:04:59		<b>makro_corr_4223_h</b> Sample 4	Healthy		<b>Healthy</b> 	







2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:05:27		<b>makro_corr_4223_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:05:53		<b>makro_corr_4226_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:06:21		<b>makro_corr_4226_i</b> Sample 4	Infected		<b>Infected</b> 	





































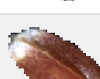







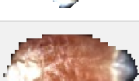



2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:06:50		<b>makro_corr_4228_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:07:18		<b>makro_corr_4228_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:07:47		<b>makro_corr_4234_h</b> Sample 4	Healthy		<b>Healthy</b> 	











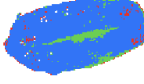


































2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:08:14		<b>makro_corr_4234_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:08:42		<b>makro_corr_4302_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:09:10		<b>makro_corr_4302_i</b> Sample 4	Infected		<b>Infected</b> 	

















































2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:09:38		<b>makro_corr_4303_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:10:06		<b>makro_corr_4303_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:10:34		<b>makro_corr_4310_h</b> Sample 4	Healthy		<b>Healthy</b> 	

















































































































2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:11:03		<b>makro_corr_4310_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:11:32		<b>makro_corr_4319_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:12:03		<b>makro_corr_4319_i</b> Sample 4	Infected		<b>Infected</b> 	

2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:12:33		<b>makro_corr_4339_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:13:02		<b>makro_corr_4339_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:13:32		<b>makro_corr_4340_h</b> Sample 4	Healthy		<b>Healthy</b> 	

2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:14:02		<b>makro_corr_4340_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:14:31		<b>makro_corr_4423_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:15:02		<b>makro_corr_4423_i</b> Sample 4	Infected		<b>Infected</b> 	

2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:15:32		<b>makro_corr_4428_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:16:03		<b>makro_corr_4428_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:21:24		<b>makro_corr_4437_h</b> Sample 4	Healthy		<b>Healthy</b> 	

2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:21:54		<b>makro_corr_4437_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:25		<b>makro_corr_4505_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:22:55		<b>makro_corr_4505_i</b> Sample 4	Infected		<b>Infected</b> 	

2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:24		<b>makro_corr_4525_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:23:55		<b>makro_corr_4525_i</b> Sample 4	Infected		<b>Infected</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 1	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 2	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 3	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:23		<b>makro_corr_4532_h</b> Sample 4	Healthy		<b>Healthy</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 1	Infected		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 2	Infected		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 3	Infected		<b>Infected</b> 	
2023-03-14 11:24:54		<b>makro_corr_4532_i</b> Sample 4	Infected		<b>Infected</b> 	



**Norges miljø- og biovitenskapelige universitet**  
Noregs miljø- og biovitenskapelige universitet  
Norwegian University of Life Sciences

Postboks 5003  
NO-1432 Ås  
Norway