

Article

Hyperspectral Remote Sensing for Early Detection of Wheat Leaf Rust Caused by *Puccinia triticina*

Anton Terentev ^{1,*}, Vladimir Badenko ^{2,*}, Ekaterina Shaydayuk ¹, Dmitriy Emelyanov ¹, Danila Eremenko ², Dmitriy Klabukov ², Alexander Fedotov ² and Viktor Dolzhenko ¹

¹ All-Russian Institute of Plant Protection, 196608 Saint Petersburg, Russia

² Advanced Digital Technologies, Peter the Great St. Petersburg Polytechnic University, 195251 Saint Petersburg, Russia

* Correspondence: admin@vizr.spb.ru (A.T.); badenko_vl@spbstu.ru (V.B.)

Abstract: Early crop disease detection is one of the most important tasks in plant protection. The purpose of this work was to evaluate the early wheat leaf rust detection possibility using hyperspectral remote sensing. The first task of the study was to choose tools for processing and analyze hyperspectral remote sensing data. The second task was to analyze the wheat leaf biochemical profile by chromatographic and spectrophotometric methods. The third task was to discuss a possible relationship between hyperspectral remote sensing data and the results from the wheat leaves, biochemical profile analysis. The work used an interdisciplinary approach, including hyperspectral remote sensing and data processing methods, as well as spectrophotometric and chromatographic methods. As a result, (1) the VIS-NIR spectrometry data analysis showed a high correlation with the hyperspectral remote sensing data; (2) the most important wavebands for disease identification were revealed (502, 466, 598, 718, 534, 766, 694, 650, 866, 602, 858 nm). An early disease detection accuracy of 97–100% was achieved from fourth dai (day/s after inoculation) using SVM.

Keywords: hyperspectral remote sensing; wheat leaf rust; *Puccinia triticina*; support vector machines; early plant disease detection; VIS-NIR spectroscopy; leaf pigments; biochemical profile



Citation: Terentev, A.; Badenko, V.; Shaydayuk, E.; Emelyanov, D.; Eremenko, D.; Klabukov, D.; Fedotov, A.; Dolzhenko, V. Hyperspectral Remote Sensing for Early Detection of Wheat Leaf Rust Caused by *Puccinia triticina*. *Agriculture* **2023**, *13*, 1186. <https://doi.org/10.3390/agriculture13061186>

Academic Editor: Feng Yang

Received: 20 April 2023

Revised: 30 May 2023

Accepted: 31 May 2023

Published: 2 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Environmental and economic problems, world economy globalization, and world population growth, are significant incentives for the development of new agricultural techniques in order to achieve food security. Precision farming has become one of the most actively developing areas of agriculture in the last two decades [1]. One of the most important areas of precision agriculture is plant disease monitoring and identification. Modern approaches include the use of various sensors that can be installed on various platforms, such as UAVs and satellites [2].

Wheat leaf rust, caused by *Puccinia triticina* Eriks., is an important disease that affects wheat production worldwide [3]. The disease causes big annual losses due to its more frequent and widespread occurrence, although it is usually less damaging than stem rust caused by *Puccinia graminis* Pers., or yellow rust caused by *Puccinia striiformis* var. *tritici* Westend [4]. Despite a decrease in the importance of *P. triticina*, noted in 2000, this pathogen can still be a source of epiphytotics capable of damaging up to 35% of the yield [5].

Hyperspectral remote sensing is one of the most promising methods for monitoring and diagnosing plant diseases [6,7]. This method allows early plant disease detection at the first disease development stages, before visible symptoms appear [8,9]. This may allow to take timely and balanced measures for pest control, excluding the excessive use of chemical pesticides [10,11].

The study of wheat leaf rust using “hyperspectral remote sensing was carried out in works of Ashourloo et al. (2014) [12] and Bohnenkamp et al. (2019) [13]. Ashourloo et al.

used different SVIs to detect the disease at different severity levels. The NBNDVI, NDVI, GI, ARI, and PRI indices showed a high-overall, disease-detection accuracy. However, these indices showed up to 50% classification errors at early stages of the disease [12]. Bohnenkamp et al. used pure fungal spore spectra, as reference, to detect the symptoms of wheat leaf rust. The model for rust detection was efficiently applied on the pixel scale, but the exact percentage of correct classification was not specified [13]. A comparative analysis of the results of both studies is provided in Table 1.

Table 1. A comparative analysis of the previous studies on the topic.

Author and Reference	Sensor Used	Sensor Type	Spectral Range	Algorithm Used	Results	Early Detection
Ashourloo et al. [12]	ASD Fieldspec 4 pro	whisk-broom camera	350–2500 nm.	NBNDVI, NDVI, PRI, GI, and RVSI	>70% accuracy	No
Bohnenkamp et al. [13]	ImSpector PFD V10E	snapshot camera	400–1000 nm.	Least-square factorization	Not mentioned	No

Both of these studies were taken into account. However, the study of Ashourloo et al. had too high of an error rate at early stages of the disease. The study of Bohnenkamp et al. involved the usage of pure fungal spore spectra, as reference, which means high-disease severity. Consequently, our study was based on the methodology of Gold et al. (2020) [14], supplemented by a number of innovations.

The authors also took into account the features of detecting various wheat diseases described in recent papers using hyperspectral remote sensing [15–17]. Based on these works, the principles of the experimental methodology and the choice of tools for analyzing the data obtained were formulated. The purpose of this work was to evaluate early wheat leaf rust detection using hyperspectral remote sensing. The first task of the study was to propose tools for processing and analyzing hyperspectral remote sensing data. The second task was to analyze wheat leaf biochemical profile by chromatographic and spectrophotometric methods. The third task was to discuss a possible relationship between hyperspectral remote sensing data and the results from the wheat leaf biochemical profile analysis. Finally, to evaluate the proposed, machine learning-based processing method for early wheat leaf rust detection.

2. Materials and Methods

2.1. Plant Material Preparation

The grains of wheat (*Triticum aestivum*) of ‘Sudarynya’ cultivar, susceptible to the wheat leaf rust, were sown into 152 pots (85 × 85 × 100 mm), in the amount of 12 grains in each pot. The pots were divided into 2 equal groups, which were used as the control and experiment, respectively.

The suspension of *P. triticina* urediniospores from the collection of the All-Russian Institute of Plant Protection in concentration 10⁴ spores/mL was used for the inoculation. The seven days old wheat seedlings were inoculated under laboratory conditions and then placed into a dark humid chamber with a 100% humidity for 10–12 h. After that, the plants were transferred into a phytotron and were grown at the temperature of 20–23 °C with the photoperiod of 14 h per day at 15,000–20,000 lux (202.5–270 PPFD (μmol m⁻² s⁻¹)), during the next 10 days.

The plant material was cut daily at 12 PM and was immediately used for hyperspectral imaging or was immediately stored for metabolomic study starting from fourth dai (day/s after inoculation). Only the first leaves were used for imaging due to the fact that the inoculation was carried out at the stage of the first leaf. The plant material storage for metabolomic study was carried out at a temperature of −80 °C using a Binder UF V 700 freezer.

The plant material preparation was made in the form of two independent repetitions carried out under identical conditions. There was a one-week interval between repetitions. The disease development was assessed by classical phytopathological methods.

2.2. Hyperspectral Camera Setup

A hyperspectral snapshot camera Ultris 20 (Cubert GmbH, Ulm, Germany) was used in the studies. The camera was operating in the range from 450 to 850 nm in the mode of 100 channels with the sampling interval of 4 nm. The hyperspectral image resolution was 410×410 pixels.

The imaging was carried out in a room isolated from natural light. The camera was installed horizontally on a tripod at a height of 0.5 m above the studied wheat leave samples. The samples were illuminated with two 500-watt halogen lamps directed at the sample at the angle of 45 degrees. A white background was used for sample shooting. The shooting area was 14 by 20 cm. The design of the installation was based on the studies [15–17], and is shown in Figure 1.

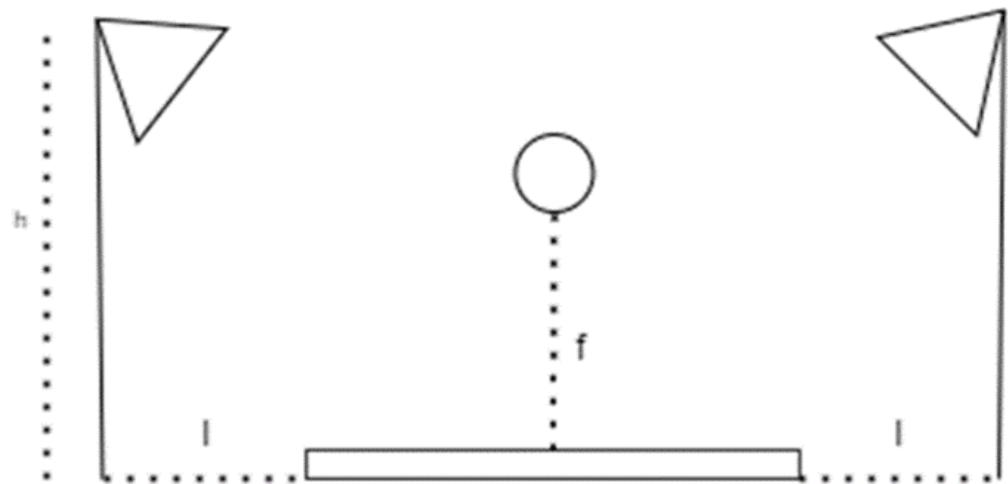


Figure 1. The shooting installation. $F = 53$ cm; $L = 80$ cm; $H = 160$ cm.

2.3. Image Acquisition

Black and white calibration, distance calibration, and PC data transfer were performed using the Cubert-Pilot software provided by the manufacturer. The information was uploaded for further processing in the Multi-Channel TIFF format with following parameters: 106 channels, 16 bits per channel. The data analysis was performed using the support vector machine (SVM) algorithm.

The images were taken from the fourth to the seventh day after inoculation every day at 12.00 A.M. Ten leaves were cut from random plants for each image. Each daily set (the control and the experiment) contained 12 images. Thus, each daily set contained 240 leave images in total, with a 960 leave images in the whole experiment, 480 healthy, and 480 diseased leaves, respectively.

2.4. Gas Chromatography-Mass Spectrometry GC-MS Analysis

The 500 mg of wheatgrass leaves (12 leaves) were ground in liquid nitrogen in a mortar with a pestle. Crude extract was obtained by methanol (80%) extraction at 40°C from these drags. The extracts analysis was done with GC-MS.

The residue was separated by centrifugation at $10,000 \times g$ for 10 min. The supernatant in a volume of 10 mL was evaporated with a vacuum-rotary evaporator at 40°C .

The compounds were identified based on their relative retention time and by comparison of their mass spectra with data from the mass-spectra library. The analysis was performed by an Agilent 5860 chromatograph using Agilent ChemStation E.02.02.1431

software (Agilent Technologies Inc., Santa Clara, CA, USA). The separation was provided with a capillary column 30 m long, 0.25 mm in diameter, and stationary phase film (95% dimethylpolyoxane, 5% diphenyl), at a thickness of 0.1 μm . The following conditions were used: helium flow rate, 1 mL/min⁻¹ and evaporator temperature, 230 °C at a flow-split ratio of 1:20. The temperature conditions of the column thermostat were the following: initial temperature of 70 °C, increased by 6 °C/min⁻¹ up to 340 °C. The peaks were recorded by an Agilent 5975S mass selective detector (Agilent Technologies, USA) in the total ion recording mode with a frequency of 1.8 scans per second. Electron impact ionization was performed at 70 V, an ion source temperature of 230 °C.

The retention-index determination was performed using a calibration based on standard, saturated hydrocarbons. Quantitative interpretations of metabolite profiles were carried out over total ion current peak areas using the internal standard method by UniChrom software (www.unichrom.com, accessed on 1 January 2023). Standard curves were defined from standards containing polyols (inositol, sorbitol, galactinol, etc.) and sugars (trehalose, sucrose, glucose, fructose, etc.).

2.5. VIS-NIR Spectroscopy

Photosynthetic pigments were extracted from leaf samples with 100% acetone and the pigment contents were calculated using the formulae described in [18]. The work was carried out using Nano-500 spectrophotometer (Hangzhou Allsheng Instruments Co., Ltd., Hangzhou, China).

3. Results

3.1. Data Classification Results

Data pre-processing. The main goal of data pre-processing was to find the pixels referring to the plant material on the hyperspectral images. It was made in two steps.

On the first step we selected the preliminary areas of interest by black and white representation using threshold clipping with selected constants [19]. On the second step, we made additional shadow filtration using NDVI threshold. This step was necessary because shadows were similar to the areas of interest in the black and white representation.

Thus, the plant material samples were separated from the background and could be used for further processing.

Data analysis. The separated plant material data was processed as follows:

- (a) **Hyperspectral curves analysis.** The first stage of data analysis was plotting hyperspectral curves [20,21]. The results obtained demonstrated similarity with the results of consimilar studies on the topic [20,22,23]. Hyperspectral curves were typical for plant objects. It was concluded that hyperspectral imaging was carried out correctly and its results could be used for further research. The spectral curves are shown in Figure 2.

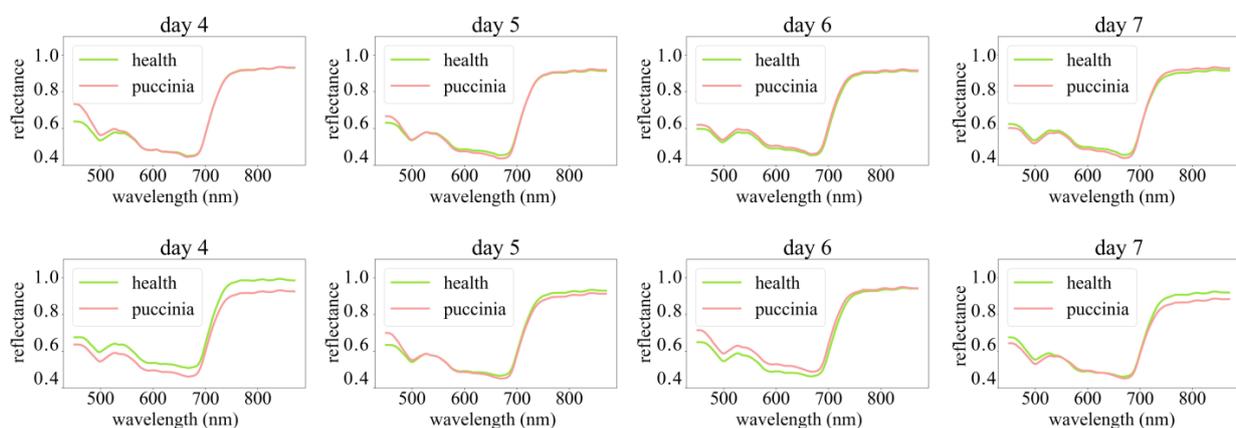


Figure 2. Averaged hyperspectral curves from two experiments.

- (b) Pixel distributions analysis. The second stage of data analysis was the search for correlated-value cluster presence and complex intensity distributions. Since hyperspectral images are inherently similar to conventional images and are a set of two-dimensional matrices, the data were analyzed in terms of correlated-value cluster presence and complex intensity distributions before classification methods choosing. The distributions were analyzed for all available wavelengths in the range of 440–870 nm. Visual hyperspectral images analysis did not demonstrate significant differences in pixel distributions, in which the control group would differ from wheat leaf rust inoculated group until 7 dai. Based on this observation, the use of neural network models to build a classification of hyperspectral images was not preferable for early disease detection. An example of wavelength matrices distributions for the 440 nm wavelength is shown in Figure 3.

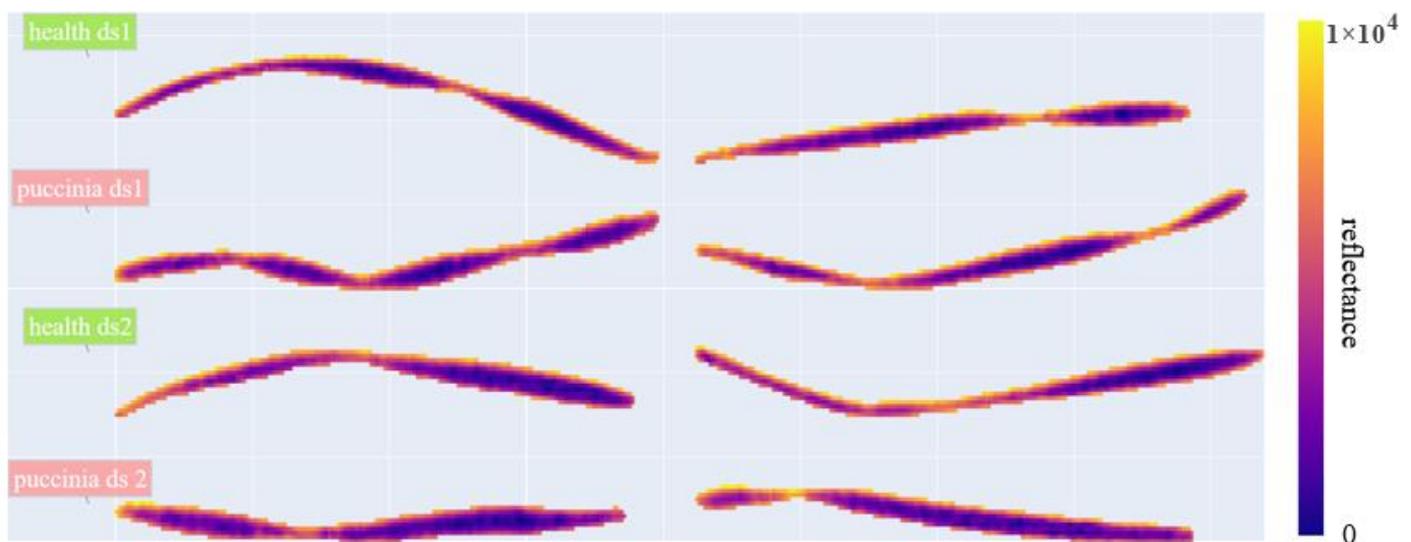


Figure 3. An example of intensity distributions in wheat hyperspectral images for the 440 nm wavelength.

- (c) Feature generation. Next, an analysis of class separability (control or experiment data) was made based on various features that could be extracted from hyperspectral images. For each available hyperspectral range, two groups of features have been calculated. The first group of features are mean values of pixel for each range. The second group are textural features for each range [24]. Homogeneity, contrast, dissimilarity and entropy have been chosen as texture features based on results from [20,25]. The average value usage for each range as a predictor was made due to an assumption that the reflectivity of diseased and healthy plants is generally different. The use of textural features was based on an assumption that control or experiment data can be different in terms of heterogeneity of their structure in certain ranges.
- (d) Preliminary analysis of separability of classes in the attribute space. To display classes in the feature space, the t-Distributed Stochastic Neighbor Embedding algorithm (t-SNE) presented in [26] was used. The algorithm is a modification of the Stochastic Neighbor-Embedding (SNE) algorithm presented in [27]. Both algorithms make it possible to map a multidimensional space onto a space of smaller dimensions, for example, a two-dimensional one. The difference of the t-SNE algorithm is that it uses a different cost function, which allows to simplify the original SNE optimization problem. T-SNE representations for different feature spaces is shown in Figure 4.

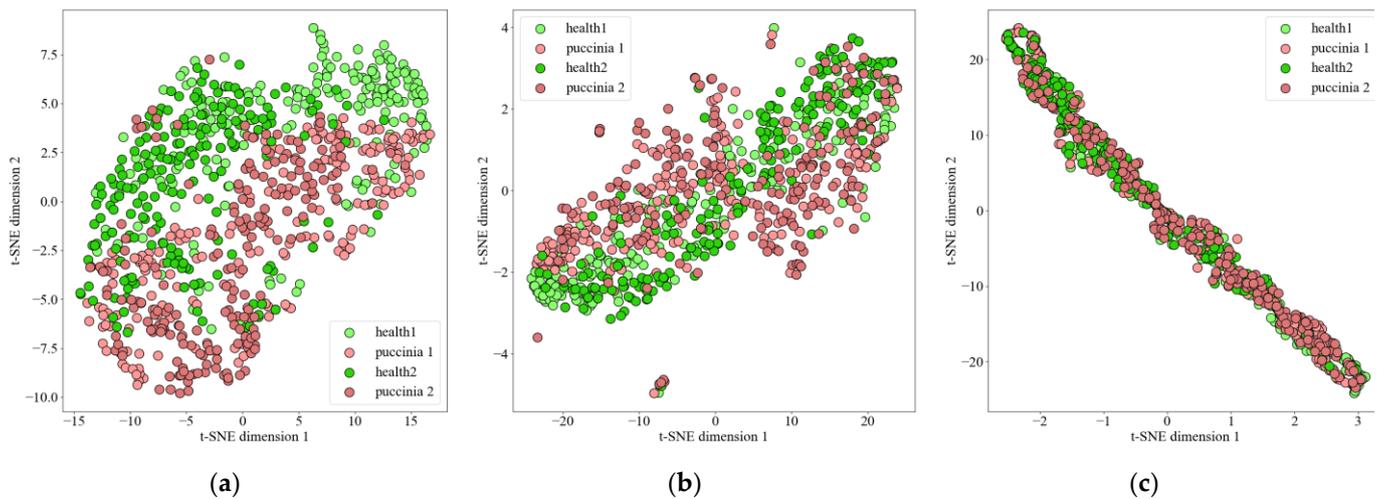


Figure 4. The classes distribution in the multidimensional space of features for images, presented in two-dimensional space using the t-SNE algorithm for cases where the features are (a) average, (b) textural features over pixel intensities in two-dimensional hyperspectral images of each channel, and (c) vegetation indices. The graphs demonstrate good separability of classes in feature spaces in cases where channel averages and texture features are used as features. The feature space built on the basis of vegetation indexes demonstrates worse separability.

Classification method, training scenario and testing. The support vector machine (SVM) method [28] and its implementation [29] were chosen as the classification method.

The SVM training was made under two scenarios. In each of the scenarios, the dataset from one of the experiments was used for training and testing. The dataset from the other experiment was used for the results evaluation. Such model testing is the most reliable and allows accuracy extrapolation assessing. Each SVM model was optimized using the grid-search algorithm on a grid of linear and non-linear kernels with a different cost of constraints (C) and sigma (σ) parameters. The experiment scheme is shown in Figure 5.

The “overall accuracy” and the “kappa score” metrics were used to evaluate the accuracy. Their values were calculated by the following formulae:

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{kappa} = \frac{2 * (TP * TN - FN * FP)}{(TP + FP) * (TN + FP) + (TP + FN) * (TN + FN)}$$

where TP is the number of true positive classifications, FP is the number of false positive classifications, TN is the number of true negative classifications, and FN is the number of false negative classifications.

In a support vector machine, the importance of channels can be obtained by evaluating absolute normalized values of the model. Such coefficients are built using the channel statistics as predictors using a linear kernel function. Since the best result was obtained using the linear SVM model, which utilized the channel means as predictors, the channel importance was evaluated based on their weights.

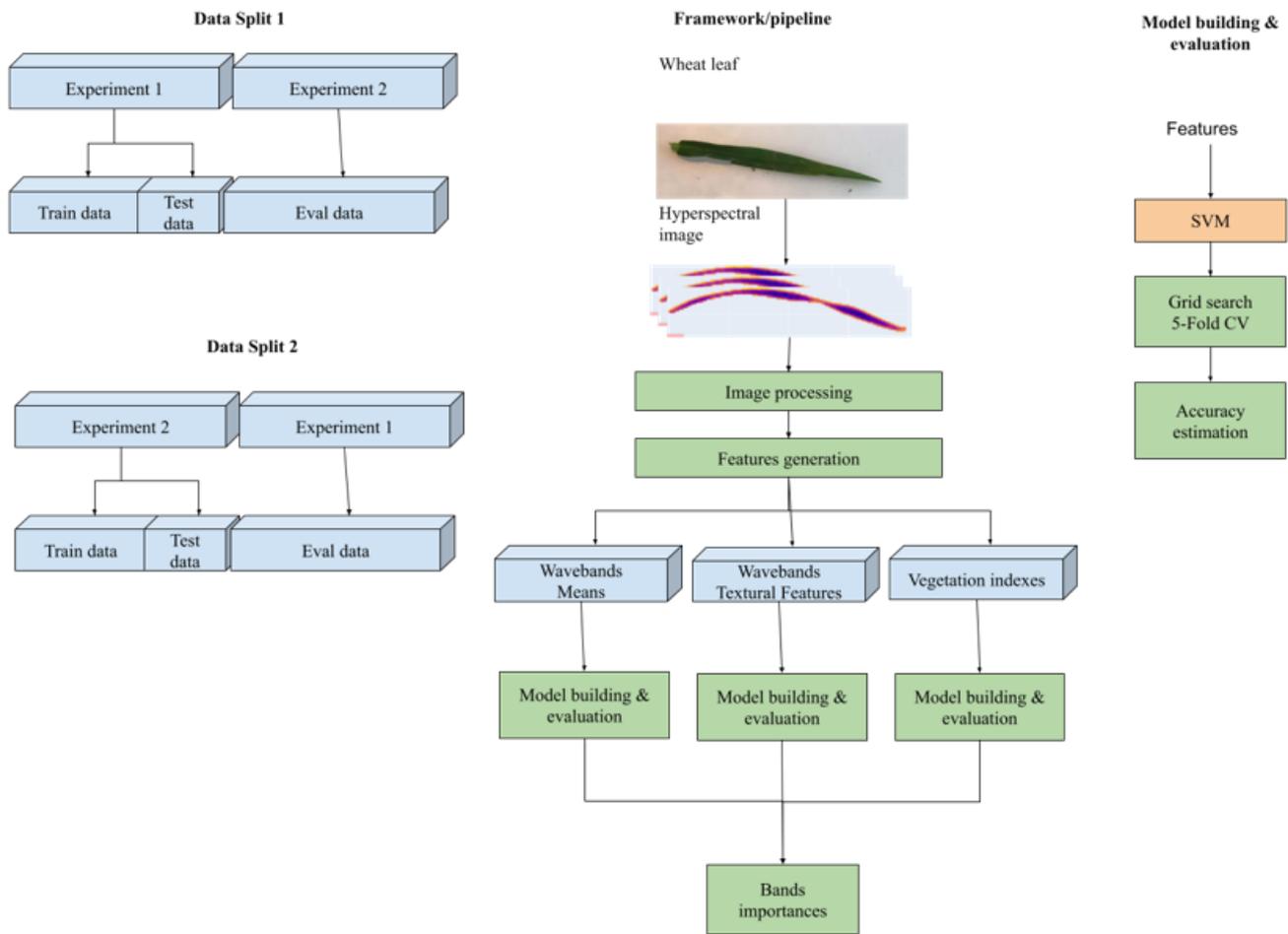


Figure 5. The experiment scheme.

To extract the smallest required set of features, a method was developed for combining importance vectors in two experiments. Each importance vector used was preliminarily multiplied by the estimated accuracy of the constructed model, after which, the combined importance was calculated as the sum of the feature importance in the two experiments divided by the difference between them.

$$imp_{i,joined} = \frac{imp_{i,exp1} * kappa_{evalexp1} + imp_{i,exp2} * kappa_{evalexp2}}{|imp_{i,exp1} * kappa_{evalexp1} - imp_{i,exp2} * kappa_{evalexp2}|}$$

$imp_{i,exp1}$ —importance of i -th feature received from SVM-model fitted on the first data split

$imp_{i,exp2}$ —importance of i -th feature received from SVM-model fitted on the second data split

$kappa_{evalexp1}$ —overall kappa score calculated on eval dataset of first data split

$kappa_{evalexp2}$ —overall kappa score calculated on eval dataset of second data split

$imp_{i,joined}$ —importance of i -th feature suitable for estimation of minimal sufficient set of features for both datasets

Hyperspectral images classification results. The hyperspectral images classification results are shown in Table 2.

Table 2. The hyperspectral images classification results.

Dataset	Model	Parameters	Features	Training		Testing		Evaluation	
				OA%	Kappa	OA%	Kappa	OA%	Kappa
set 1	SVM	C = 1 kernel = linear	MEAN	100.0	1.00	99.0	0.98	98.0	0.97
		C = 1000 kernel = linear	TF	100.0	1.00	98.0	0.97	97.0	0.94
		C = 1000 gamma = 0.1 kernel = rbf	INDEXES	100.0	0.99	98.0	0.97	0.94	0.88
		C = 1000 gamma = 0.0001 kernel = linear	MEAN + TF	100.0	1.00	100.0	1.00	98.0	0.96
set 2	SVM	C = 1 kernel = linear	MEAN	100.0	1.00	100.0	1.00	98.0	0.97
		C = 100 gamma = 0.001 kernel = rbf	TF	100.0	1.00	98.0	0.95	95.0	0.90
		C = 1000 kernel = linear	INDEXES	99.0	0.99	98.0	0.97	94.0	0.88
		C = 100 gamma = 0.0001 kernel = linear	MEAN + TF	99.0	0.99	98.0	0.95	93.0	0.86

C is a regularization parameter. The strength of the regularization is inversely proportional to C. MEAN is the average for each channel, TF are textural features for each channel, INDEXES are vegetation indices.

The table shows the classifiers built on the following groups of features. The best result was shown by the SVM model trained on the average values for each channel as input features. It can be concluded that the most effective approach was to analyze the reflectivity of diseased and healthy plants in general. The use of indices demonstrated the worst average result. The texture features usage showed good results for the first data split (kappa_eval = 0.94 vs. kappa_eval = 0.97), but was significantly lower for the model built on the channel averages in the second data split (kappa_eval = 0.90 vs. kappa_eval = 0.97). Additionally, combining the two best approaches MEAN and TF did not provide better results than the MEAN approach alone.

A mutual portability between experiments was observed. The smallest necessary set of wavebands necessary to build a classifier suitable for both experiments was selected. The most important wavebands ordered by importance are 502, 466, 598, 718, 534, 766, 694, 650, 866, 602, 858 nm. Figure 6 shows the channel importance vectors for two experiments (on the left), and the combined importance vector for two experiments (on the right).

Figure 7 shows the dependence of the model accuracy on the most important wavebands number used for data split one, on the left, and for data split two, on the right.

3.2. Metabolomic Analysis Results

GC-MS data evaluation. The metabolomic data evaluation showed that the entire spectrum of primary and secondary wheat metabolites, such as sugars, amino acids, organic acids, etc., changed significantly from the first hours of wheat leaf rust development. PCA (principal component analysis) has shown a significant increase in the content of gamma-aminobutyric acid, alpha and beta alanine, a decrease in the content of carbohydrates, and an increase in the concentration of glycosylated phenols (Figure 8).

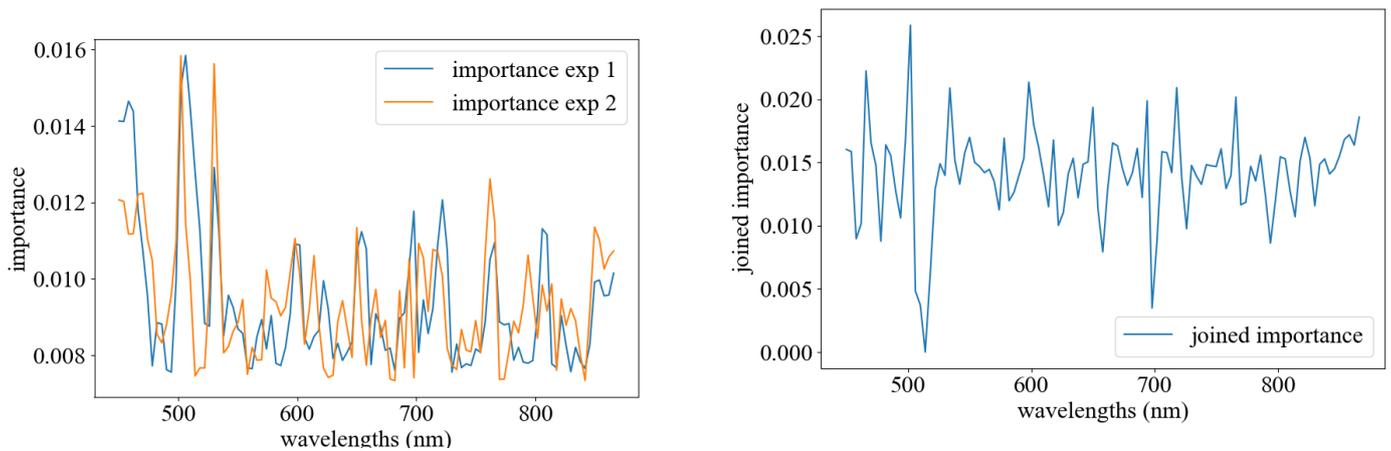


Figure 6. An example of intensity distributions for the wavelength of 440 nm.

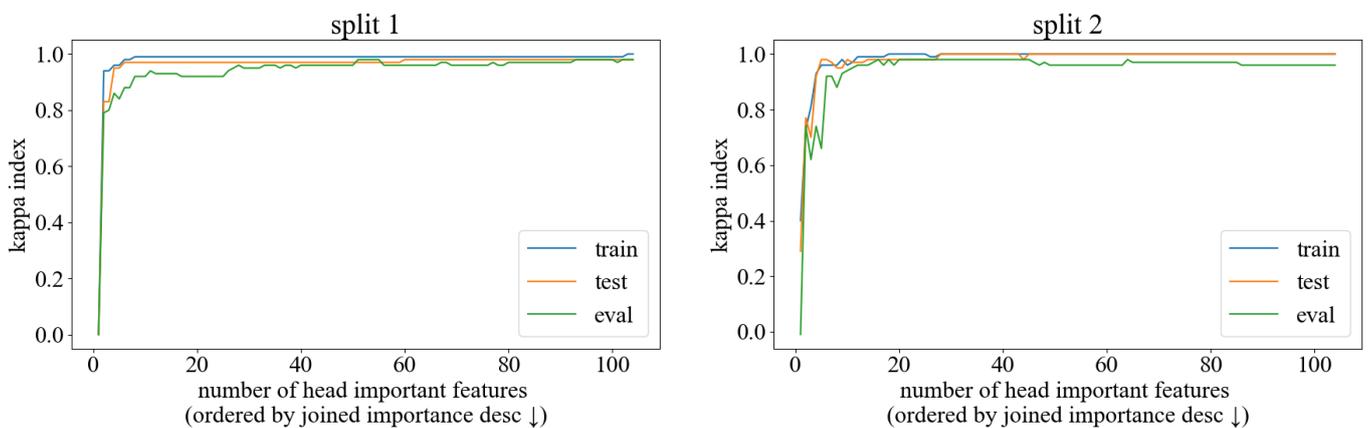


Figure 7. The figure shows that the model built on the basis of the 11 most important channels in terms of the imp_{joined} metric, demonstrates a classification accuracy of 97–100%, while the model built on the basis of all 104 channels demonstrates an accuracy of 95–99%.

The highest amounts of amino acids were observed in non-inoculated wheat plants. During the development of the disease (4 dai), their number significantly decreased, which, apparently, is associated with active biosynthesis of phenolamides, wheat phytoalexins [30,31], in response to pathogen inoculation. Additionally, during the development of the disease, the number and variety of di- and monosaccharides increased, and specialized metabolites appeared (retention times 36.43 and 41.92).

It was impossible to detect most of these compounds with the Cubert Ultris 20 camera, as the camera's operating range is VIS-NIR (450–850 nm), while these compounds' detection ranges are UV (100–400 nm) and SWIR (1100–2500 nm) [32,33]. Thus, metabolomics data had an extremely limited use to refine hyperspectral VIS-NIR remote sensing data for wheat leaf rust detection.

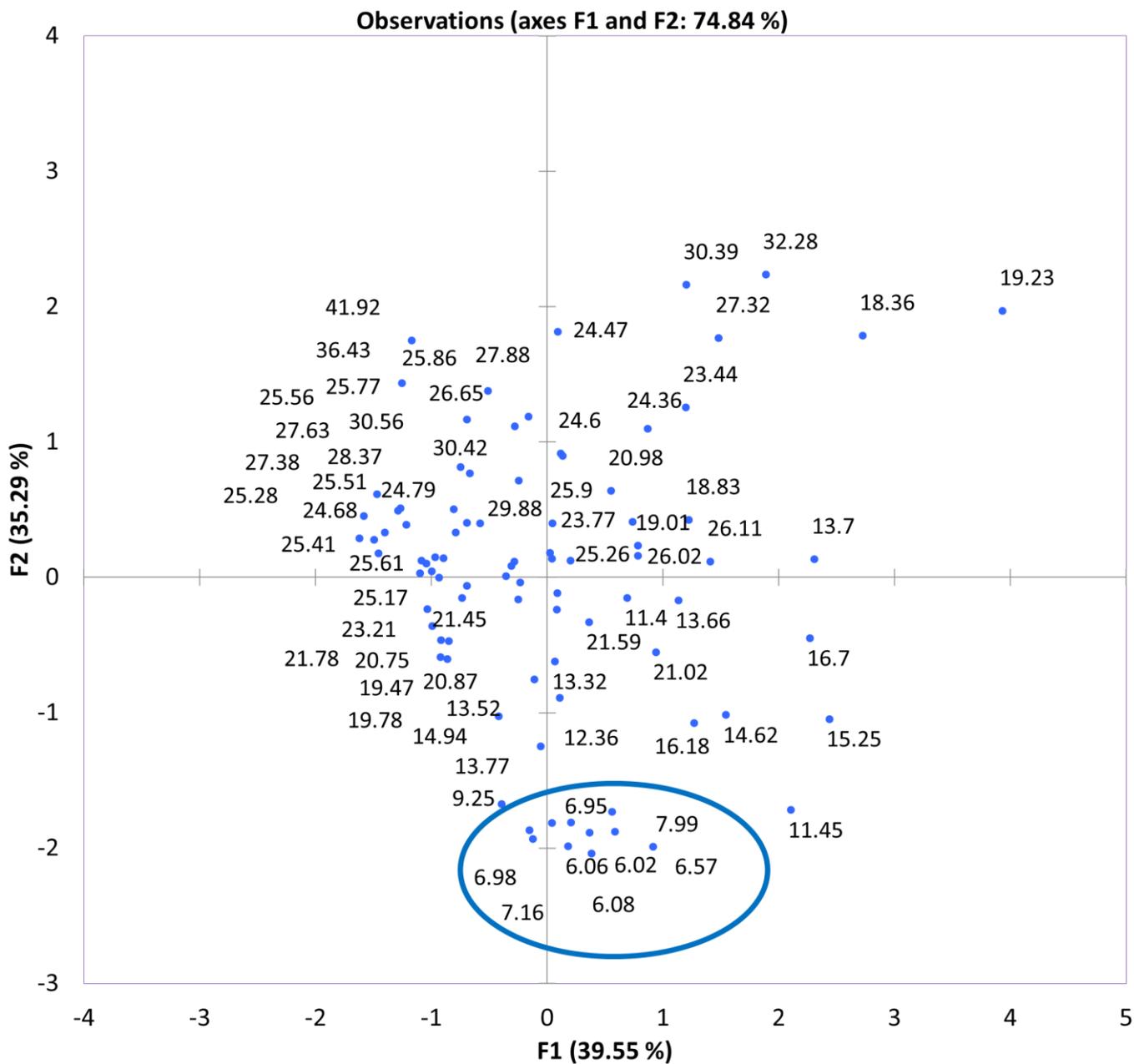


Figure 8. Principal component analysis of wheat metabolome on 4 dai (components by retention time).

Spectroscopy data evaluation. The trend in changes of *chlorophyll a* and *b* and total carotenoids concentration is similar for both hyperspectral (Figure 9) and spectrometry (Figure 10) data. The following formulae for the pigment specific ratio were used for the calculation when using the hyperspectral data:

$$PSSR \text{ for Chlorophyll } a = \frac{R800}{R675}$$

$$PSSR \text{ for Chlorophyll } b = \frac{R800}{R650}$$

$$PSSR \text{ for Carotenoids} = \frac{R800}{R500}$$

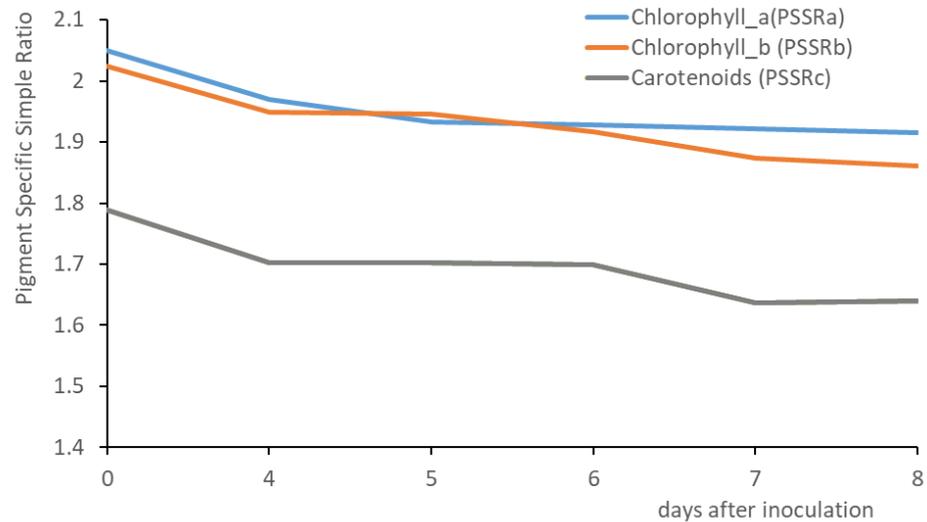


Figure 9. Chlorophylls (a, b) and total carotenoids simple ratio dynamics in wheat leaf rust inoculated leaves measured by hyperspectral, remote sensing, Pigment Specific Simple Ratio by Blackburn (1998); $LSD_{0.05} = 0.05$ [34].

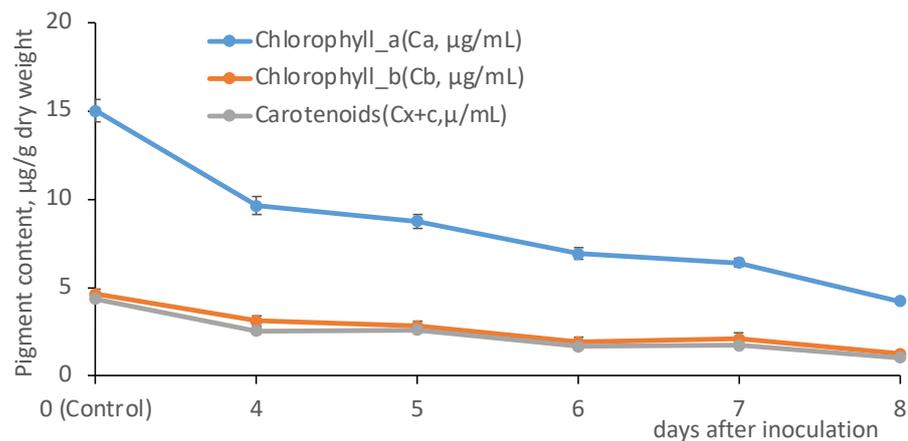


Figure 10. Chlorophylls (a, b) and total carotenoids level dynamics in wheat leaf rust-inoculated leaves measured via VIS-NIR spectrometry by Lichtenthaler and Buschmann (2001); $LSD_{0.05} = 1.1$ [18].

The following formulae for pigment contents were used for the calculation when using the spectrometry data:

$$\text{Chlorophyll a (ca, } \mu\text{g/mL)} = 11.24 A_{662} - 2.04 A_{645}$$

$$\text{Chlorophyll b (cb, } \mu\text{g/mL)} = 20.13 A_{645} - 4.19 A_{662}$$

$$\text{Carotenoids (cx + c, } \mu\text{g/mL)} = (1000 A_{470} - 1.90 ca - 63.14 cb) / 214$$

$$\text{Chlorophyll a (mg/g dw)} = ca * V/W$$

$$\text{Chlorophyll b (mg/g dw)} = cb * V/W$$

$$\text{Carotenoids (mg/g dw)} = cx + c * V/W$$

where V = volume of solvent (5 mL), and W = dry weight of leaf sample [18].

It follows that wheat leaf rust detection with hyperspectral remote sensing is based on pigment concentration changes. At the same time, the method sensitivity makes it possible

to detect the disease at early stages before visible symptoms appear. Thus, hyperspectral remote sensing, which can detect wheat leaf rust at 4 dai, has a significant advantage over classical phytopathological methods.

4. Discussion

Based on the analysis of hyperspectral remote sensing for early plant disease detection, a number of assumptions were made about the biological experiment design and its successful repetition. Of those applicable to this experiment, we found it necessary to take into account a number of important factors. Firstly, the spectral portraits' differences can be caused by various abiotic and biotic stress factors [7,35]. Secondly, the presence of a mixed infection can also be important [7,8,36]. Finally, we made an assumption that it may be important to understand what biochemical changes occur inside a diseased plant, in what wavelength range they manifest themselves, and how they can affect the spectral portrait [37–39]. We believed that hyperspectral imaging data are determined by the biochemical processes occurring in plants from the first minutes of disease development.

Therefore, we decided to conduct the experiment in laboratory conditions, in the absence of any factors other than the disease under study itself. The hyperspectral data acquisition was made under controlled-lighting conditions. This made it possible to obtain two independent sets of data from two similar experiments.

There is a sufficient number of articles in the field of early plant disease detection using hyperspectral remote sensing [6–9,35,36]. From these articles, it follows that machine learning methods for early plant disease detection generally require significant initial data [7,8]. Different authors used different analytical tools for processing such data [40–43].

In the study by Alisaak et al. (2018), a differentiation accuracy of 76% between non-inoculated and *Fusarium* head blight infected plants using SVM was achieved. Specim ImSpector V10E and ImSpector N25E spectrographs operating in 400–1000 nm and 1000–2500 nm ranges were used in the study [44]. Mahlein et al. (2019) achieved early detection (3rd dai) of *Fusarium graminearum* inoculated wheat, using SVM with an accuracy of 78% via the same equipment [21]. Finally, the same results were obtained by Huang et al. (2020) who reached a detection accuracy of 75% with the SVM method, studying the same disease with the same sensors [45]. Despite the significant biological differences in the disease development caused by *Fusarium graminearum*, and the herein studied *Puccinia triticina*, they do not play a key role in the algorithm learning, so an algorithm with the proven efficiency was used.

The review by Golhani et al. (2018) highlighted the current state of Imaging and non imaging hyperspectral data for early disease detection with different types of artificial neural networks (ANN) [42]. The authors have shown incredible capabilities in adapting the neural networks for disease detection purposes using hyperspectral data. The review presents materials from which, it follows, that it is possible to use different types of neural networks to determine plant diseases with a high accuracy (>97%). Morellos et al. (2020) used machine learning methods to estimate the existence of the tomato chlorosis virus and viral load in the tomato leaves [46]. The multilayer perceptron with automated, relevance determination (MLP-ARD) classifier showed the best results with an overall accuracy of 92.1%.

SVM is a supervised learning method, in which the optimization function is to maximize the distance between the separating plane and the elements of the training sample closest to this plane in feature space. The SVM method allows to make non-linear separation of classes using different kernel functions extending feature space. SVM demonstrates good accuracy in solving the problem of classifying plant diseases from hyperspectral data [20,47]. Additionally, the method choice was made due to its interpretability when constructed without the use of nonlinear kernels and the small size of the collected dataset.

The convolutional neural network (CNN) approach could not be used in this work, despite this method being successfully applied in earlier studies [48,49]. As it was already mentioned at the data analysis section, there was no significant difference between the control group and the diseased plant hyperspectral images in terms of pixel distributions.

Neural networks are promising for extracting hard-to-distinguish visible features [50,51]. We decided that it is necessary to use a higher-resolution camera to extract such patterns. Since the Cubert Ultris 20 camera has a resolution of 410×410 pixels, such features may not be distinguishable.

Various decision tree-based methods were not used in this study because of the known issue that a good accuracy of such methods is achieved by building ensembles of data-independent solvers, which requires a much bigger training set [52,53]. Thus, the support vector machine (SVM) method and its implementation were chosen as the classification method.

The results of the metabolomic study of biochemical changes occurring in the plant did not give a significant result. The data obtained from the Cubert Ultris 20 hyperspectral camera, which works in the actual range of 450–950 nm, cover only an insignificant part of the compounds identified as a result of the metabolome comparative analysis of the diseased and healthy wheat. Leaf pigments make up only a part of the metabolome, and thus, at this stage, the GC-MS data analysis cannot be compared with the VIS-NIR hyperspectral remote sensing data. Based on the literature analysis, it was found that the compounds identified by the metabolomic analysis, whose concentration changes significantly during the course of the disease, generally can be detected in the range of 1100–2500 nm. Therefore, when using a camera with a range of 450–950 nm, one cannot rely on metabolomics for the verification of the disease under study. Accordingly, the use of cameras with a working range that includes the range of 1100–2500 nm may allow more accurate characterization of the development of the disease using metabolomics tools. In the existing articles, this possibility was not considered, but it should be noted that, in general, the higher verification percentage of the diseased and healthy plants was significantly higher for devices operating in the 400–2500 nm range [8]. Recently, the studies on UV-hyperspectral remote sensing data usage, that rely on metabolomics data approval, have appeared [39].

The VIS-NIR spectrometry data analysis showed a high correlation with the hyperspectral remote sensing data. The signs of wheat leaf rust development were identified by comparing the content of the three main pigments (chlorophylls a and b, and total carotenoids) in the control and experiment groups, according to the previously published methods [12]. The detected changes level in the chlorophylls and carotenoids' concentration turned out to be lower using hyperspectral remote sensing than using VIS-NIR spectrometry (Figures 9 and 10). However, this sensitivity was sufficient to reliably classify the diseased and healthy specimens at 4 dai and beyond.

Thus, the developed method relies on the SVM-based analysis of the proven changes in the pigments' composition during wheat leaf rust development. We believe that further data accumulation on this disease's development is necessary, both in field and laboratory conditions. In the future, this will make it possible to make a quick conclusion about the disease's spread and intensity degree from a small sample. This may allow timely wheat leaf rust control.

The results of this study showed that hyperspectral remote sensing is a promising method for early wheat leaf rust detection. The authors plan to carry out further work to verify the obtained data in the field. From a technical and physical point of view, it is necessary to take into account the model of sunlight propagation during such an experiment, since hyperspectral remote sensing is a passive method that depends on ambient light. Other technical problems may include improper use of the equipment. Proper calibration of the hyperspectral sensor or camera is essential for proper data acquisition. This calibration depends on the temperature of the hyperspectral sensor; thus, the equipment must be calibrated after a continuous period of operation, in order to offset the effect of its heating.

In our opinion, further research directions may also include a comparison of wheat leaf rust with other fungal pathogens. It would be interesting to compare hyperspectral images obtained as a result of plants exposure to other biotrophic and necrotrophic fungi. The use of hyperspectral remote sensing data in the wavelengths of 1100–2500, can also be

useful for developing new algorithms for early plant disease detection. The new generation of high-resolution hyperspectral camera usage, may allow more correct classification of inoculated plants at an even earlier time.

5. Conclusions

The purpose of this work was to evaluate early wheat leaf rust detection possibility using hyperspectral remote sensing. The results obtained showed that hyperspectral remote sensing and the proposed, machine learning-based processing method is the promising method for early wheat leaf rust detection. The chosen SVM classification method reached the data classification accuracy of 97–100% for the model built on the 11 most important wavebands 502, 466, 598, 718, 534, 766, 694, 650, 866, 602, 858 nm. The VIS-NIR spectrometry data analysis showed a high correlation with the VIS-NIR hyperspectral remote sensing data. The GC-MS data analysis could not be compared with the VIS-NIR hyperspectral remote sensing data due to the fact that leaf pigments make up only a part of the metabolome. Thus, the discussed method for early wheat leaf rust detection is based on SVM classification of the proven changes in the pigments' composition during the disease development, measured by hyperspectral remote sensing.

Supplementary Materials: The following supporting information can be downloaded at: <https://disk.yandex.ru/d/mpmijVmolslIFQ>.

Author Contributions: Conceptualization, A.T., V.D., V.B. and A.F.; methodology, A.T., E.S., V.B. and V.D.; software, D.K. and D.E. (Danila Eremenko); validation, D.E. (Danila Eremenko), D.K. and D.E. (Dmitriy Emelyanov); formal analysis, D.E. (Danila Eremenko); investigation, A.T., E.S., D.K. and D.E. (Dmitriy Emelyanov); resources, E.S. and A.F.; data curation, A.F. and V.D.; writing—original draft preparation, A.T. and D.E. (Danila Eremenko); writing—review and editing, A.F., E.S., D.K., D.E. (Dmitriy Emelyanov) and V.D.; visualization, D.E. (Dmitriy Emelyanov); supervision, V.B. and V.D.; project administration, A.T., V.B. and V.D.; funding acquisition, A.T., V.B. and A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research is partially funded by the Ministry of Science and Higher Education of the Russian Federation as part of the World-class Research Center program: Advanced Digital Technologies (contract No. 075-15-2022-311 dated 20 April 2022).

Data Availability Statement: Data are contained within the article or Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
ARI	Anthocyanin Reflectance Index
CNN	Convolutional Neural Network
GC-MS	Gas Chromatography-Mass Spectrometry
GI	Greenness Index
LSD	Least Significant Difference
MLP-ARD	Multilayer Perceptron with Automated Relevance Determination
NBNDVI	Narrow-Band Normalized Difference Vegetation Index
NDVI	Normalized Difference Vegetation Index
PRI	Photochemical Reflectance Index
VIS-NIR	Visible/Near Infrared
SNE	Stochastic Neighbor Embedding algorithm
SVM	Support Vector Machine
SWIR	Short Wave Infrared
TIFF	Tag Image File Format
t-SNE	t-Distributed Stochastic Neighbor Embedding algorithm
UAV	Unmanned Aerial Vehicle

References

- Gebbers, R.; Adamchuk, V. Precision agriculture and food security. *Science* **2010**, *327*, 828–831. [CrossRef]
- Shafi, U.; Mumtaz, R.; García-Nieto, J.; Hassan, S.A.; Zaidi, S.A.R.; Iqbal, N. Precision Agriculture Techniques and Practices: From Considerations to Applications. *Sensors* **2019**, *19*, 3796. [CrossRef]
- Bolton, M.; Kolmer, J.; Garvin, D. Wheat leaf rust caused by *Puccinia triticina*. *Mol. Plant Patol* **2008**, *9*, 563–575. [CrossRef]
- Huerta-Espino, J.; Singh, R.P.; German, S.; McCallum, B.D.; Park, R.F.; Chen, W.Q.; Bhardwaj, S.C.; Goyeau, H. Global status of wheat leaf rust caused by *Puccinia triticina*. *Euphytica* **2011**, *179*, 143–160. [CrossRef]
- Gulyaeva, E.; Shaydayuk, E.; Gannibal, P. Leaf Rust Resistance Genes in Wheat Cultivars Registered in Russia and Their Influence on Adaptation Processes in Pathogen Populations. *Agriculture* **2021**, *11*, 319. [CrossRef]
- Mahlein, A.K.; Kuska, M.T.; Thomas, S.; Bohnenkamp, D.; Alisaac, E.; Behmann, J.; Wahabzada, M.; Kersting, K. Plant disease detection by hyperspectral imaging: From the lab to the field. *Adv. Anim. Biosci.* **2017**, *8*, 238–243. [CrossRef]
- Khan, A.; Vibhute, A.D.; Mali, S.; Pati, C.H. A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications. *Ecol. Inform.* **2022**, *69*, 101678. [CrossRef]
- Terentev, A.; Dolzhenko, V.; Fedotov, A.; Eremenko, D. Current State of Hyperspectral Remote Sensing for Early Plant Disease Detection: A Review. *Sensors* **2022**, *22*, 757. [CrossRef] [PubMed]
- Wan, L.; Li, H.; Li, C.; Wang, A.; Yang, Y.; Wang, P. Hyperspectral Sensing of Plant Diseases: Principle and Methods. *Agronomy* **2022**, *12*, 1451. [CrossRef]
- Finger, R. No pesticide-free Switzerland. *Nat. Plants* **2021**, *7*, 1324–1325. [CrossRef]
- Moysiadis, V.; Sarigiannidis, P.; Vitsas, V.; Khelifi, A. Smart farming in Europe. *Comput. Sci. Rev.* **2021**, *39*, 100345. [CrossRef]
- Ashourloo, D.; Mobasheri, M.R.; Huete, A. Evaluating the Effect of Different Wheat Rust Disease Symptoms on Vegetation Indices Using Hyperspectral Measurements. *Remote Sens.* **2014**, *6*, 5107–5123. [CrossRef]
- Bohnenkamp, D.; Kuska, M.T.; Mahlein, A.K.; Behmann, J. Hyperspectral signal decomposition and symptom detection of wheat rust disease at the leaf scale using pure fungal spore spectra as reference. *Plant Pathol.* **2019**, *68*, 1188–1195. [CrossRef]
- Gold, K.M.; Townsend, P.A.; Chlus, A.; Herrmann, I.; Couture, J.J.; Larson, E.R.; Gevens, A.J. Hyperspectral Measurements Enable Pre-Symptomatic Detection and Differentiation of Contrasting Physiological Effects of Late Blight and Early Blight in Potato. *Remote Sens.* **2020**, *12*, 286. [CrossRef]
- Chen, J.; Saimi, A.; Zhang, M.; Liu, Q.; Ma, Z. Epidemic of Wheat Stripe Rust Detected by Hyperspectral Remote Sensing and Its Potential Correlation with Soil Nitrogen during Latent Period. *Life* **2022**, *12*, 1377. [CrossRef]
- Huang, L.; Liu, Y.; Huang, W.; Dong, Y.; Ma, H.; Wu, K.; Guo, A. Combining Random Forest and XGBoost Methods in Detecting Early and Mid-Term Winter Wheat Stripe Rust Using Canopy Level Hyperspectral Measurements. *Agriculture* **2022**, *12*, 74. [CrossRef]
- Liu, Y.; Zhang, Y.; Jiang, D.; Zhang, Z.; Chang, Q. Quantitative Assessment of Apple Mosaic Disease Severity Based on Hyperspectral Images and Chlorophyll Content. *Remote Sens.* **2023**, *15*, 2202. [CrossRef]
- Lichtenthaler, H.K.; Buschmann, C. Chlorophylls and Carotenoids: Measurement and Characterization by UV-VIS Spectroscopy. *Curr. Protoc. Food Anal. Chem.* **2001**, *1*, F4.3.1–F4.3.8. [CrossRef]
- Horprasert, T.; Harwood, D.; Davis, L.S. A statistical approach for real-time robust background subtraction and shadow detection. *IEEE ICCV* **1999**, *99*, 1–19.
- Guo, A.; Huang, W.; Ye, H.; Dong, Y.; Ma, H.; Ren, Y.; Ruan, C. Identification of Wheat Yellow Rust Using Spectral and Texture Features of Hyperspectral Images. *Remote Sens.* **2020**, *12*, 1419. [CrossRef]
- Mahlein, A.-K.; Alisaac, E.; Al Masri, A.; Behmann, J.; Dehne, H.-W.; Oerke, E.-C. Comparison and Combination of Thermal, Fluorescence, and Hyperspectral Imaging for Monitoring Fusarium Head Blight of Wheat on Spikelet Scale. *Sensors* **2019**, *19*, 2281. [CrossRef]
- Guo, A.; Huang, W.; Dong, Y.; Ye, H.; Ma, H.; Liu, B.; Wu, W.; Ren, Y.; Ruan, C.; Geng, Y. Wheat Yellow Rust Detection Using UAV-Based Hyperspectral Technology. *Remote Sens.* **2021**, *13*, 123. [CrossRef]
- Zhang, X.; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A Deep Learning-Based Approach for Automated Yellow Rust Disease Detection from High-Resolution Hyperspectral UAV Images. *Remote Sens.* **2019**, *11*, 1554. [CrossRef]
- Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [CrossRef]
- Xuan, G.; Li, Q.; Shao, Y.; Shi, Y. Early Diagnosis and Pathogenesis Monitoring of Wheat Powdery Mildew Caused by *Blumeria Graminis* Using Hyperspectral Imaging. *Comput. Electron. Agric.* **2022**, *197*, 106921. [CrossRef]
- Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
- Hinton, G.E.; Roweis, S.T. Stochastic neighbor embedding. *Adv. Neural Inf. Process. Syst.* **2002**, *15*, 857–864.
- Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
- Available online: <https://scikit-learn.org/stable/modules/svm.html> (accessed on 1 January 2023).
- Ube, N.; Harada, D.; Katsuyama, Y.; Osaki-Oka, K.; Tonooka, T.; Ueno, K.; Taketa, S.; Ishihara, A. Identification of phenylamide phytoalexins and characterization of inducible phenylamide metabolism in wheat. *Phytochemistry* **2019**, *167*, 112098. [CrossRef]
- Ube, N.; Yabuta, Y.; Tohnooka, T.; Ueno, K.; Taketa, S.; Ishihara, A. Biosynthesis of Phenylamide Phytoalexins in Patho-gen-Infected Barley. *Int. J. Mol. Sci.* **2019**, *20*, 5541. [CrossRef]

32. Curran, P.J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278. [[CrossRef](#)]
33. Jacquemoud, S.; Ustin, S. Leaf optical properties in different wavelength domains. In *Leaf Optical Properties*; Jacquemoud, S., Ustin, S., Eds.; Cambridge University Press: Cambridge, UK, 2019; pp. 124–169.
34. Blackburn, G.A. Quantifying Chlorophylls and Carotenoids at Leaf and Canopy Scales. *Remote Sens. Environ.* **1998**, *66*, 273–285. [[CrossRef](#)]
35. Mahlein, A.K.; Kuska, M.T.; Behmann, J.; Polder, G.; Walter, A. Hyperspectral sensors and imaging technologies in phytopathology: State of the art. *Annu. Rev. Phytopathol.* **2018**, *56*, 535–558. [[CrossRef](#)]
36. Bouguettaya, A.; Zarzour, H.; Kechida, A.; Taberkit, A.M. A survey on deep learning-based identification of plant and crop diseases from UAV-based aerial images. *Cluster Comput.* **2023**, *26*, 1297–1317. [[CrossRef](#)]
37. Chen, F.; Ma, R.; Chen, X.-L. Advances of Metabolomics in Fungal Pathogen–Plant Interactions. *Metabolites* **2019**, *9*, 169. [[CrossRef](#)]
38. Castro-Moretti, F.R.; Gentzel, I.N.; Mackey, D.; Alonso, A.P. Metabolomics as an Emerging Tool for the Study of Plant–Pathogen Interactions. *Metabolites* **2020**, *10*, 52. [[CrossRef](#)]
39. Brugger, A.; Yamati, F.I.; Barreto, A.; Paulus, S.; Schramowski, P.; Kersting, K.; Steiner, U.; Neugart, S.; Mahlein, A.-K. Hyperspectral Imaging in the UV Range Allows for Differentiation of Sugar Beet Diseases Based on Changes in Secondary Plant Metabolites. *Phytopathology* **2023**, *113*, 44–54. [[CrossRef](#)]
40. Behmann, J.; Mahlein, A.-K.; Rumpf, T.; Römer, C.; Plümer, L. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precis. Agric.* **2015**, *16*, 239–260. [[CrossRef](#)]
41. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
42. Golhani, K.; Balasundram, S.K.; Vadamalai, G.; Pradhan, B. A review of neural networks in plant disease detection using hyperspectral data. *Inf. Process. Agric.* **2018**, *5*, 354–371. [[CrossRef](#)]
43. Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad, M. Recent Advances in Image Processing Techniques for Automated Leaf Pest and Disease Recognition—A Review. *Inf. Process. Agric.* **2020**, *8*, 27–51. [[CrossRef](#)]
44. Alisaac, E.; Behmann, J.; Kuska, M.T.; Dehne, H.-W.; Mahlein, A.-K. Hyperspectral quantification of wheat resistance to Fusarium head blight: Comparison of two Fusarium species. *Eur. J. Plant Pathol.* **2018**, *152*, 869–884. [[CrossRef](#)]
45. Huang, L.; Zhang, H.; Ruan, C.; Huang, W.; Hu, T.; Zhao, J. Detection of scab in wheat ears using in situ hyperspectral data and support vector machine optimized by genetic algorithm. *Int. J. Agric. Biol. Eng.* **2020**, *13*, 182–188. [[CrossRef](#)]
46. Morellos, A.; Tziotziou, G.; Orfanidou, C.; Pantazi, X.E.; Sarantaris, C.; Maliogka, V.; Alexandridis, T.K.; Moshou, D. Non-Destructive Early Detection and Quantitative Severity Stage Classification of Tomato Chlorosis Virus (ToCV) Infection in Young Tomato Plants Using Vis–NIR Spectroscopy. *Remote Sens.* **2020**, *12*, 1920. [[CrossRef](#)]
47. Baek, I.; Kim, M.S.; Cho, B.K.; Mo, C.; Barnaby, J.Y.; McClung, A.M.; Oh, M. Selection of Optimal Hyperspectral Wavebands for Detection of Discolored, Diseased Rice Seeds. *Appl. Sci.* **2019**, *9*, 1027. [[CrossRef](#)]
48. Nagasubramanian, K.; Jones, S.; Singh, A.K.; Sarkar, S.; Singh, A.; Ganapathysubramanian, B. Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods* **2019**, *15*, 98. [[CrossRef](#)]
49. Polder, G.; Blok, P.M.; de Villiers, H.A.C.; van der Wolf, J.M.; Kamp, J. Potato Virus Y Detection in Seed Potatoes Using Deep Learning on Hyperspectral Images. *Front. Plant Sci.* **2019**, *10*, 209. [[CrossRef](#)]
50. Gomez Villa, A.; Salazar, A.; Vargas, F. Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecol. Inform.* **2017**, *41*, 24–32. [[CrossRef](#)]
51. Buetti-Dinh, A.; Galli, V.; Bellenberg, S.; Ilie, O.; Herold, M.; Christel, S.; Boretska, M.; Pivkin, I.V.; Wilmes, P.; Sand, W.; et al. Deep Neural Networks Outperform Human Expert’s Capacity in Characterizing Bioleaching Bacterial Biofilm Composition. *Biotechnol. Rep.* **2019**, *22*, e00321. [[CrossRef](#)]
52. Ke, G.L.; Meng, Q.; Finley, T.; Wang, T.F.; Chen, W.; Ma, W.D.; Ye, Q.W.; Liu, T.Y. LightGBM: A highly efficient gradient boosting decision tree. In Proceedings of the 31st Annual Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 3146–3154.
53. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018; pp. 6639–6649.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.