

Article

A Gated Recurrent Units (GRU)-Based Model for Early Detection of Soybean Sudden Death Syndrome through Time-Series Satellite Imagery

Luning Bi ¹, Guiping Hu ^{1,*}, Muhammad Mohsin Raza ², Yuba Kandel ², Leonor Leandro ² and Daren Mueller ²

¹ Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, IA 50011, USA; luningbi@iastate.edu

² Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA 50011, USA; mohsinuaf@live.com (M.M.R.); ykandel@iastate.edu (Y.K.); lleandro@iastate.edu (L.L.); dsmuelle@iastate.edu (D.M.)

* Correspondence: gphu@iastate.edu; Tel.: +1-515-294-8638

Received: 7 October 2020; Accepted: 2 November 2020; Published: 4 November 2020



Abstract: In general, early detection and timely management of plant diseases are essential for reducing yield loss. Traditional manual inspection of fields is often time-consuming and laborious. Automated imaging techniques have recently been successfully applied to detect plant diseases. However, these methods mostly focus on the current state of the crop. This paper proposes a gated recurrent unit (GRU)-based model to predict soybean sudden death syndrome (SDS) disease development. To detect SDS at a quadrat level, the proposed method uses satellite images collected from PlanetScope as the training set. The pixel image data include the spectral bands of red, green, blue and near-infrared (NIR). Data collected during the 2016 and 2017 soybean-growing seasons were analyzed. Instead of using individual static imagery, the GRU-based model converts the original imagery into time-series data. SDS predictions were made on different data scenarios and the results were compared with fully connected deep neural network (FCDNN) and XGBoost methods. The overall test accuracy of classifying healthy and diseased quadrates in all methods was above 76%. The test accuracy of the FCDNN and XGBoost were 76.3–85.5% and 80.6–89.2%, respectively, while the test accuracy of the GRU-based model was 82.5–90.4%. The calculation results show that the proposed method can improve the detection accuracy by up to 7% with time-series imagery. Thus, the proposed method has the potential to predict SDS at a future time.

Keywords: soybean disease; sudden death syndrome; gated recurrent unit; remote sensing; satellite imagery; disease detection

1. Introduction

Soybean (*Glycine max* L. Merrill) diseases can have a significant impact on production and profits [1]. During the years from 2015 to 2019, soybean diseases were responsible for losses of around 8.99% of the production potential in the U.S., which equates to an average of USD 3.8 billion annually [2]. Sudden death syndrome (SDS) is one of the most damaging soybean diseases found throughout most of the soybean production area in the United States. SDS is caused by a soilborne fungus *Fusarium virguliforme* (*Fv*) that causes root rot and foliar symptoms that typically become visible during reproductive stages [3]. Visual assessment of SDS requires intensive crop scouting that is time-consuming and labor-intensive. Therefore, an automated method for the detection of SDS is necessary.

Timely detection and management of plant diseases are essential for reducing yield loss. Remote sensing technology has proven to be a practical, noninvasive tool for monitoring crop growth and assessing plant health [4,5]. Sensing instruments can record radiation in various parts of the electromagnetic spectrum, including ultraviolet, visible, near-infrared (NIR) and thermal infrared, to name a few [6]. Healthy and diseased plant canopies absorb and reflect incident sunlight differently due to changes in leaf and canopy morphology and chemical constituents [7,8]. These changes can alter the optical spectra, such as a decrease in canopy reflectance in the near-infrared band and an increase in reflectance in the red band [7]. Some widely used methods include thermography [9–11], fluorescence measurements [12–14] and hyperspectral techniques [15–17].

In the past few years, several studies have been conducted for the detection of plant diseases [18–20]. For example, Durmus et al. [18] used red, green and blue (RGB) cameras for disease-detection on tomato leaves. In another study, Gold et al. [19] used lab-based spectroscopy to detect and differentiate late and early blight diseases on potato leaves. Specific to SDS, Bajwa et al. [21] and Herrmann et al. [3] used handheld and tractor-mounted sensors, respectively, for SDS detection. These successful applications help ease the demand for expert resources and reduce the human errors in the visual assessment and, eventually, management of plant diseases.

However, most of these methods are near-sensing techniques that focus on individual plants. In practice, it is more reasonable and efficient to diagnose plant diseases such as SDS from the quadrat level. There are several options for collecting imagery for the detection of plant diseases at the quadrat level, including unmanned aerial vehicles (UAV) [22,23], tractor-mounted tools [3] and satellite imagery [24]. Satellite imagery is typically less expensive, covers a wide ground swath and can provide temporal flexibility because the fields can be continuously monitored non-destructively. The frequency of image collection is determined by the number of satellites that pass that field [25]. Although the imagery from UAVs and tractor-mounted tools is often higher resolution at the quadrat level, these tools are cumbersome for farmers to operate and maintain in a commercial system and often lack spatial information. Satellite imagery, on the other hand, contains spatial information, comes preprocessed and continuously improves in resolution.

In addition to quality data, efficient and accurate analysis of the sensor data is essential for accurate detection of plant diseases. For the analysis of visible sensing information, several machine learning methods have been used. Common methods that have been used in plant disease diagnosis using images are convolutional neural networks (CNNs) and random forest. Different from full connected neural networks, CNNs have two special layer types. The convolutional layers extract features from the input images. The pooling layers reduce the dimensionality of the features. Dhakate et al. [26] used a CNN for the recognition of pomegranate diseases with 90% overall accuracy. Ferentinos [27] developed CNN models to classify the healthy and diseased plants with 99.5% success. Polders et al. [28] used fully convolutional neural networks to detect potato virus in seed potatoes. However, in our case, the resolution of satellite imagery was $3\text{ m} \times 3\text{ m}$, which means the satellite imagery of each generalized quadrat only contained several pixels. CNNs or other methods are less suitable for this task. How to deal with low-pixel-level remote sensing data is important for improving classification accuracy. Random forest was also used for disease detection and classification [29]. Samajpati et al. [30] used a random forest classifier to classify different apple fruit diseases. Chaudhary et al. [31] proposed an improved random forest method, which achieved 97.8% for a multi-class groundnut disease dataset. Although random forest is popular for its easy implementation and high efficiency on large datasets, its performance is influenced by the hyperparameter choices, such as random seeds and the number of variables [32].

Most of the current plant disease identification methods use field images at a single time-point to identify the contemporary status of the disease. The use of temporal image sequences can help improve detection accuracy. For example, a recurrent neural network (RNN) was designed for solving the multivariate time-series prediction problem [33]. However, RNN is faced with gradient vanishing/exploding problems [34]. As an improved version of RNN, the long short-term memory

model (LSTM) is used for its successful application to natural language modeling [35]. Compared with RNN, LSTM has more gates that can control the reset of the memory and the update of the hidden states. Turkoglu et al. [36] proposed an LSTM-based CNN for the detection of apple diseases and pests, which scored 96.1%. Namin et al. [37] utilized a CNN-LSTM framework for plant classification of various genotypes as well as the prediction of plant growth and achieved an accuracy of 93%. Navarro et al. [38] proposed an LSTM method for the prediction of low temperatures in agriculture and obtained a determination coefficient (accuracy) of 99%. Although LSTM has alleviated the gradient vanishing/exploding problem of RNNs, the training speed of LSTM is much slower due to the increased number of parameters. To solve this issue, Chung et al. [39] introduced the gated recurrent unit (GRU) in 2014. Since GRU only has two gates (i.e., reset gate and update gate) and uses the hidden state to transfer information, its training speed is much faster. Jin et al. [40] used a deep neural network which combined CNN and GRU to classify wheat hyperspectral pixels and obtained an accuracy of 0.743.

The objective of this paper is to detect SDS in soybean quadrats using 3m x 3m resolution satellite imagery which contains RGB and NIR information. The GRU-based model was compared in different scenarios (i.e., percentage of diseased quadrats) to the most popular neural network and tree-based methods, namely, a fully connected deep neural network (FCDNN) and XGBoost.

This paper is organized as follows. Section 2 introduces the dataset and methods used. Section 3 includes a case study, the calculation results and comparisons of the three methods. The paper concludes with the summary, findings and future research directions in Section 4.

2. Materials and Methods

Data were collected in 2016 and 2017 from an ongoing soybean field experiment located at the Iowa State University Marsden Farm, in Boone County, Iowa (Figure 1) [29,41–43]. This study site was chosen because soybean plots in this experiment have consistently displayed a range of SDS levels since 2010 [44]. In the trial site, there were three cropping systems, 2-year, 3-year and 4-year crop rotations, which are represented by using one-hot encoding, i.e., “100”, “010” and “001”, respectively. In the 2-year cropping rotation, corn and soybean were planted in rotation with synthetic fertilizers at the conventional rate. In the 3-year cropping rotation, corn, soybean, oat and red clover were planted in rotation with composted cattle manure and synthetic fertilizers at reduced rates. In the 4-year cropping rotation, corn, soybean, oat and alfalfa were planted in rotation with composted cattle manure and synthetic fertilizers at reduced rates. More information about the experiment can be found in these studies [29,41–43]. In each year, 240 soybean quadrats (3 m wide × 1.5 m long) were marked to collect disease data. The soybean variety (Latham L2758 R2, maturity group 2.7) remained the same for all seasons.

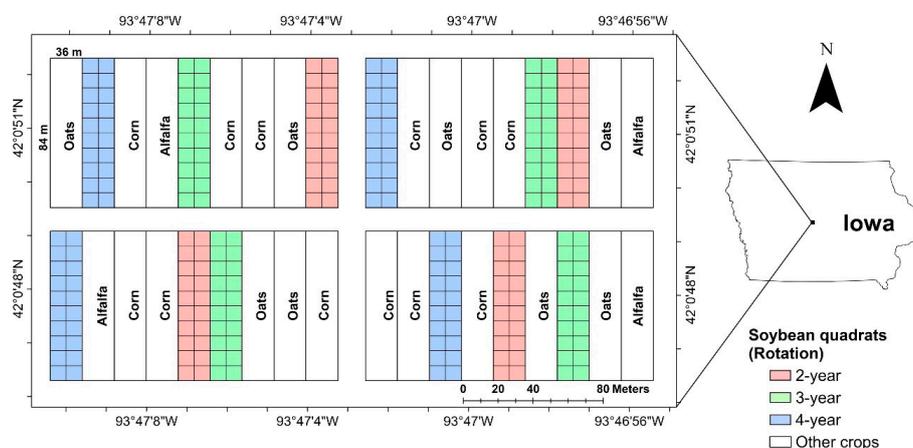


Figure 1. Experimental layout of the Marsden Farm located in Boone County, Iowa. The experiment was designed using a randomized complete block design with four blocks and each block has nine main plots. Each soybean plot was divided into 20 quadrats (8 m × 9 m, shown as square grids).

2.1. Data Processing

Figure 2 illustrates our method for data collection, data processing and analysis for the detection of SDS in this study. This methodology has been explained in detail.

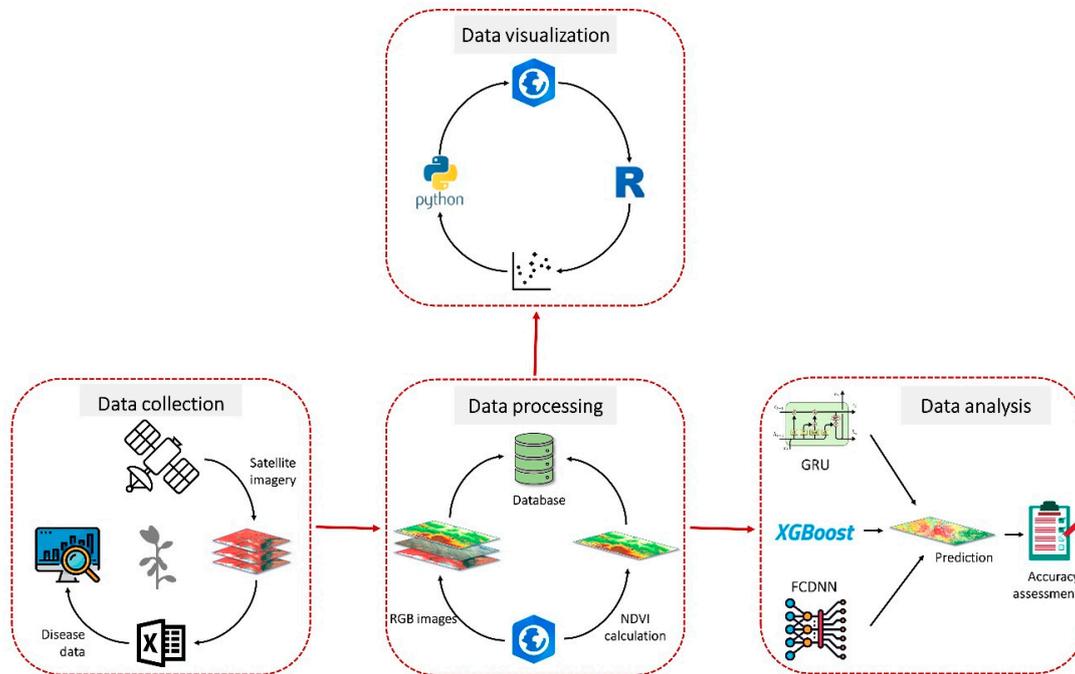


Figure 2. Flow diagram of the data collection, processing and analysis we employed in this study for sudden death syndrome (SDS) detection. We divided our methodology into four major steps including data collection, data processing, data visualization and analysis.

Satellite images were collected on 5 July 2016, 9 July 2016, 20 July 2016, 5 August 2016, 21 August 2016, 31 August 2016, 5 July 2017, 9 July 2017, 20 July 2017, 2 August 2017, 18 August 2017 and 23 August 2017. In addition to spectral information, the dataset also included ground-based crop rotation information that is an explanatory variable. The data source of satellite imagery is PlanetScope (<https://www.planet.com/>) satellite operated by Planet Labs (San Francisco, CA), a private imaging company. PlanetScope satellite imagery comes with four bands, including red (590–670 nm), green (500–590 nm), blue (455–515 nm) and NIR (780–860 nm). Soybean quadrats were generalized to large quadrats (8.6 mW × 9.1 mL) for data extraction from images and subsequent data analysis. As such, the imagery of each quadrat has 6–9 pixels.

The number of total plants and plants showing foliar symptoms of SDS was counted in each quadrat to calculate the disease incidence on a 0 to 100% scale, based on the percentage of diseased plants. The distribution of SDS incidence is shown in Figure 3. It can be observed that in 2016, SDS incidence in more than half of the 240 quadrats was less than 5% and the SDS incidence of the rest quadrats ranged from 5% to 100%. In 2017, the SDS incidence of most quadrats was below 5%. In this paper, if a quadrat had an SDS incidence above 5%, it was considered as diseased (positive); otherwise, it was considered as healthy (negative). The inspection dates used for the analysis were 27 July 2016, 5 August 2016, 22 August 2016 and 29 August 2016. Since the visual disease scores on 27 July 2016 were all zeros, all the imagery collected before that date was labeled as healthy. The human visual ratings recorded on 5 August 2016 were mapped to the imagery collected on the same date. The visual score on 22 August 2016 was mapped to the imagery collected on 21 August 2016, while the SDS rating recorded on 31 August 2016 was mapped to the imagery collected on 29 August 2016. Similarly, in 2017, the human visual ratings recorded on 17 August 2017 were mapped to the imagery collected on 18 August and the human visual ratings recorded on 24 August 2017 were mapped to the imagery collected on 23 August 2017. There was only one diseased quadrat on 5 August 2016. In 2016, all quadrats were

healthy before 5 August 2016. In 2017, all quadrats were healthy before 17 August 2017. The aim was to determine whether the quadrat had SDS or not.

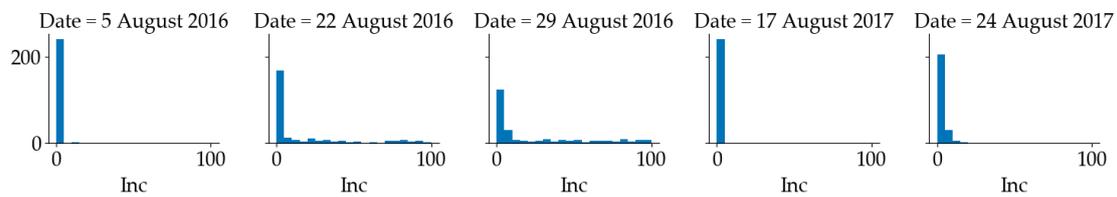


Figure 3. Distribution of sudden death syndrome (SDS) incidence in soybean quadrats (Inc.: Incidence) at Marsden Farm.

The mean and variance of the normalized RGB, normalized difference vegetation index (NDVI) and NIR information of each quadrat were used as predictors for two reasons. First, the number of pixels varied from quadrat to quadrat. Second, the SDS rating represented the entire quadrat. Some information could be lost if each pixel were used separately. The crop rotation information was used as a categorical variable encoded by one-hot method. Therefore, there were a total of 12 variables for each quadrat, which included the year, crop rotation, the mean values and the variance values of RGB and NIR bands and NDVI. The NDVI is calculated as Equation (1).

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \quad (1)$$

The box plots of RGB and NIR of diseased quadrats and healthy quadrats are shown in Figure 4. Since all quadrats were healthy on the other dates, only the results on later dates have been plotted. It can be noticed that the RGB values of diseased quadrats were lower than those of healthy quadrats, while the NIR values of diseased quadrats were greater than those of healthy quadrats. This indicates that SDS does influence light emissions of leaves. As such, one of the objectives of this paper is to use these predictors to detect the SDS-infected quadrats. To avoid mutual interference between data samples collected in a quadrat at different time points, the dataset was divided into a training and a testing dataset according to quadrat number. For the experiment in each year, we randomly selected 200 quadrats (83% of the dataset) as the training dataset and 40 quadrats (17% of the dataset) as the testing dataset of the total 240 quadrats.

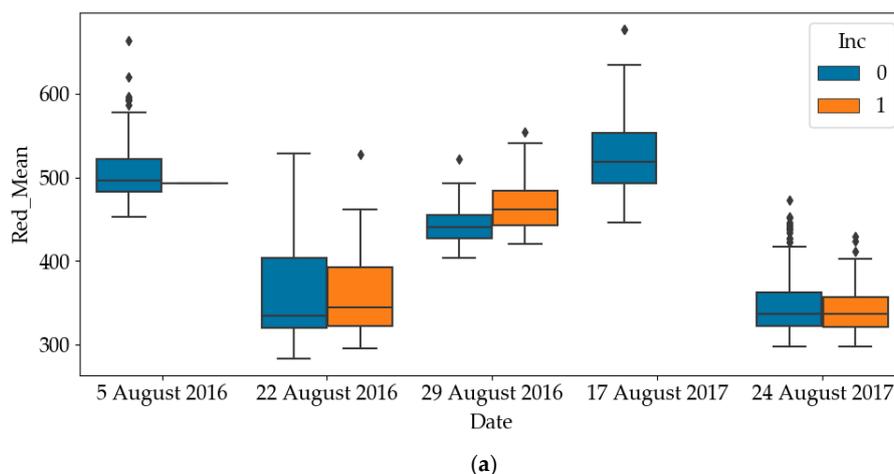
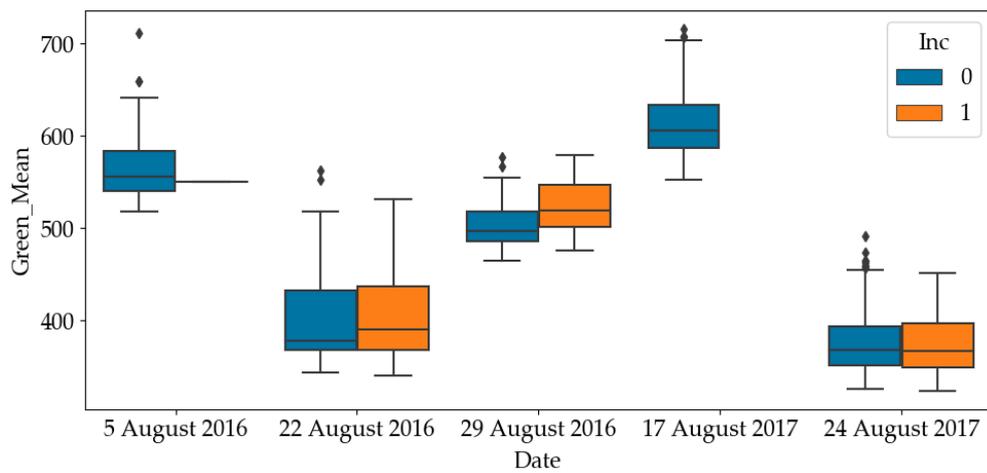
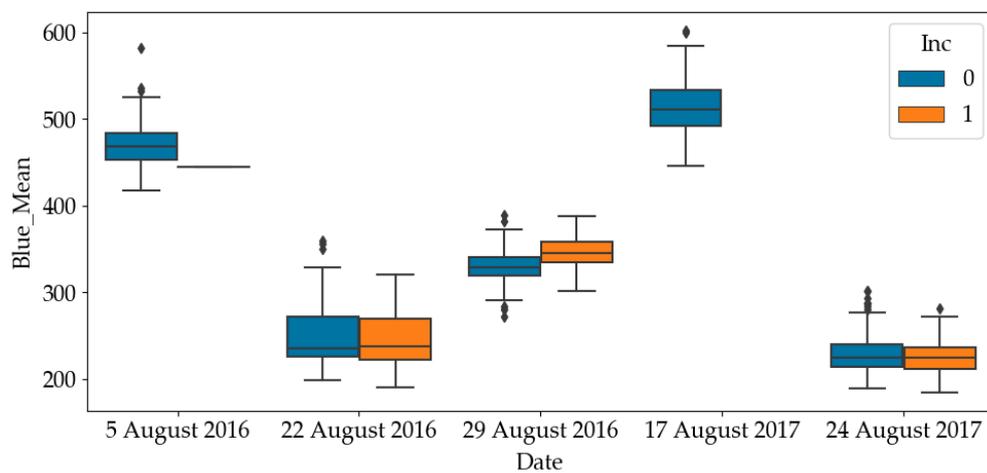


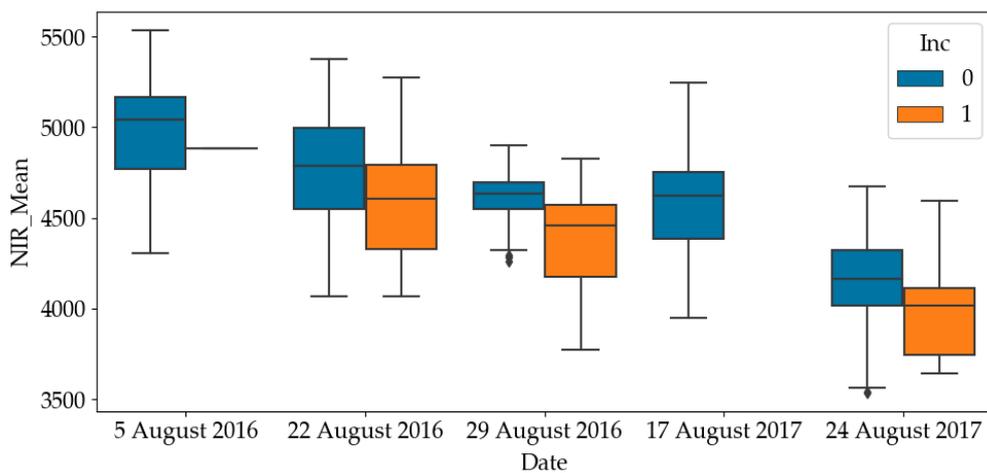
Figure 4. Cont.



(b)



(c)



(d)

Figure 4. Differences in red, green, blue and near-infrared (NIR) values between healthy and diseased quadrats (Inc=0: Healthy; Inc=1: Diseased). (a) Difference in red values between healthy and diseased quadrats; (b) Difference in green values between healthy and diseased quadrats; (c) Difference in blue values between healthy and diseased quadrats; (d) Difference in NIR values between healthy and diseased quadrats.

2.2. Measurements

Three indices, including overall accuracy, precision and recall, as calculated in Equations (2)–(4), were measured to evaluate the performance of models. The definitions of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) are shown in Table 1.

$$\text{OverallAccuracy} = (\text{TP} + \text{TN}) / \text{Total} \tag{2}$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \tag{3}$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \tag{4}$$

Table 1. Definitions of true positive (TP), false positive (FP), true negative (TN) and false negative (FN).

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

2.3. Methods

Three methods were investigated in this paper. The first is the GRU-based method. The second is XGBoost, which is the representative of tree-based methods. The third is an FCDNN, which is the most widely used deep learning method. As for comparisons, both XGBoost and the FCDNN used individual images while the GRU-based method used time-series images.

2.3.1. Gated Recurrent Unit (GRU)-Based Method

Most of the existing methods use one individual spectral measurement to predict the corresponding SDS [3]. Nevertheless, the multiple images at different times of the same quadrat may help improve the prediction accuracy of the model. Thus, a method that can handle time-series imagery is needed. Recurrent neural networks (RNNs) are suitable for non-linear time-series processing [45]. As shown in Figure 5, the RNN consists of an input layer x , a hidden layer h and an output layer y . When dealing with time-series data, the RNN can be unfolded as the right part. The output and hidden layers can be calculated according to Equation (5) and Equation (6), respectively.

$$y_t = g(s_t * w_{hy}) \tag{5}$$

$$s_t = f(x_t * w_{sx} + s_{t-1} * w_{ss}) \tag{6}$$

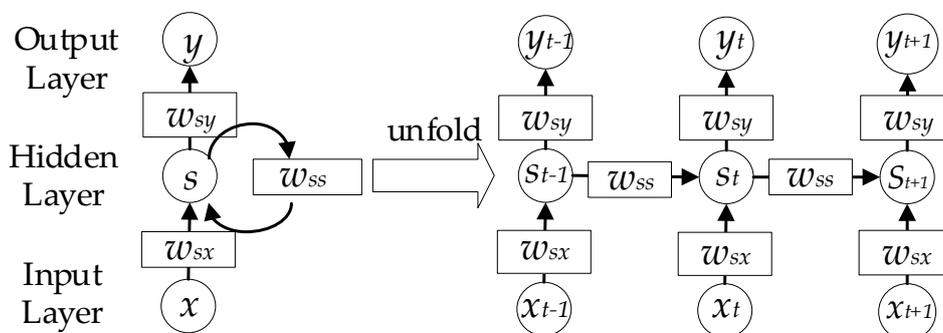


Figure 5. Structure of the recurrent neural network. At each time step, the network uses the output and internal state from the previous time step as the input of the current time step.

Despite its popularity as a universal function approximator and easy implementation, the RNN method is faced with the gradient vanishing/exploding problem. In the training process of RNNs, gradients are calculated from the output layer to the first layer of the RNN. If the gradients are smaller than 1, the gradients of the first several layers will become small through many multiplications. On the contrary, the gradients will become very large if the gradients are larger than 1. Therefore, it sometimes causes the gradients to become almost zero or very large when it reaches the first layers of RNNs. Consequently, the weights of the first layers will not get updated in the training process. Therefore, simple RNNs may not be suitable for some complex problems.

In this paper, a GRU-based method is proposed to deal with the multivariate time-series imagery data that will solve the vanishing gradient problem of a standard RNN [39]. As shown in Figure 6, based on the previous output h_{t-1} and the current input x_t , a reset gate is used to determine which part of information should be reset, as calculated in Equation (7), while an update gate is used to update the output of the GRU h_t , as calculated in Equation (8). The candidate hidden layer is calculated according to Equation (9). The current output can be obtained according to Equation (10). The gates, namely, z_t and r_t , and parameters, namely, W_z , W_r and W , of the GRU were updated in the training process.

$$z_t = \sigma(W_z \bullet [h_{t-1}, x_t]) \tag{7}$$

$$r_t = \sigma(W_r \bullet [h_{t-1}, x_t]) \tag{8}$$

$$h'_t = \tanh(W \bullet [r_t * h_{t-1}, x_t]) \tag{9}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * h'_t \tag{10}$$

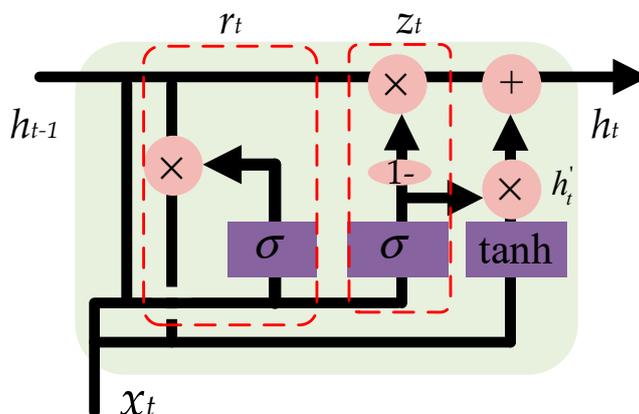


Figure 6. Structure of gated recurrent unit (GRU). The reset gate and the update gate take the previous output and the current input as the input.

2.3.2. Fully Connected Deep Neural Network (FCDNN)

FCDNN is a class of methods that use multiple layers to extract information from the input data [46]. The basic layers are a fully connected layer and an activation layer. The fully connected layer consists of multiple neurons. Each neuron in a fully connected layer connects to all neurons in the next layer. The output of a fully connected layer is calculated as Equation (11).

$$y = Wx + b \tag{11}$$

The fully connected layer can only deal with a linear problem. To add the non-linear characteristic to the model, the concept of activation layers was introduced. Some widely used activation functions

include sigmoid function, hyperbolic tangent function (Tanh) and rectified linear unit (ReLU) function. In this paper, ReLU is used as Equation (12).

$$y = \max(0, x) \quad (12)$$

The loss function is used to measure the performance of models. In this paper, binary cross-entropy loss function is used, which can be calculated as Equation (13),

$$L = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))) \quad (13)$$

where N is the total number of the samples, i is the index of the sample, y_i is the label of the i th sample and $p(y_i)$ is the predicted probability of the sample belonging to the class.

The most common training method is stochastic gradient descent which calculates the gradients and then updates the weights and biases iteratively. The final goal of training is to minimize the loss function. FCDNNs can deal with large datasets and execute feature engineering without explicit programming.

2.3.3. XGBoost

The XGBoost method is a popular tree-based method for classification tasks [47]. A decision tree consists of three parts: branches, internal nodes and leaf nodes. The internal nodes are a set of conditions that can divide the samples into different classes. The branches represent the outcome of internal nodes. The leaf nodes represent the label of the class. A decision-tree can break down a complex classification problem into a set of simpler decisions at each stage [48].

One of the most famous tree-based methods is random forest. Random forest is an ensemble method that grows many trees and outputs the class based on the results of each individual tree. Given a training set X with labels Y , it randomly selects samples with replacement of the training set and train the trees, $b_1, b_2, b_3, \dots, b_B$. After training, for sample x' , the prediction can be made by averaging the predictions of each tree or using the majority voting method. Since the bootstrap sampling method can reduce the variance of the model without increasing the bias, the model can have better generalization performance. Since each tree is trained separately, it can be implemented in parallel, which can save a lot of time. Another merit of random forest is that it can deal with high-dimensional dataset and identify the importance of each variable. Different from a random forest algorithm, which generates many trees at the same time, XGBoost is trained in an additive manner. In each iteration of XGBoost training, a new decision tree that minimizes the prediction error of previous trees will be added to the model. The prediction of a sample can be calculated by summarizing the scores of each tree. XGBoost also includes a regularization term in the loss function, which can further improve the generalization ability of the model. The specific details are given in [47].

3. Results

To validate the proposed technique, numerical calculations were conducted. The Python code can be found in the Supplementary Materials.

3.1. Model Parameters

The FCDNN consists of six fully connected layers which include 12, 8, 8, 8, 4 and 1 neurons, respectively. After each fully connected layer, a ReLU layer and a dropout layer are included. The dropout rate is 0.2. The binary cross-entropy function is used as the loss function of the model. The optimizer is stochastic gradient descent. To avoid overfitting, the dropout technique was also used for each fully connected layer. The learning rate was set as 0.008. The batch size was 100. The maximum number of iterations was set to 300.

For XGBoost, the max number of leaves was set as 40. The max depth was set as 10. The objective is the binary cross-entropy function. The maximum number of iterations was 300.

For the proposed model, two GRU layers were stacked. The first GRU layer converts each input sequence to a sequence of 50 units. Then, the second GRU layer converts the sequences of 50 units to sequences of 25 units. Then, a fully connected layer is used to transform the 25 variables to one output. The learning rate was set as 0.012. The batch size was 100. The maximum number of iterations was set to 300. All three models used the cross-entropy loss function as Equation (13).

3.2. Calculations in Different Scenarios

To determine the effectiveness of the GRU-based method in different scenarios (i.e., the percentage of diseased quadrats), four analyses have been conducted. Since both the dataset of 2016 and that of 2017 had six pairs of imagery and human visual ratings, the dates of each year were numbered from 1 to 6. All the soybean plants at time points 1 to 3 in 2016 and time points 1 to 4 in 2017 were healthy; therefore, to reduce the influence of imbalanced data, we targeted to predict SDS at time points 3 to 6 in 2016 and 5 to 6 in 2017.

The specific settings of each calculation are shown in Table 2. In calculation I, the target was to predict SDS at time points 3 to 6 in 2016 and time points 5 and 6 in 2017. In the training of the GRU-based model, the imagery collected at time points 1 and 2 was added to construct the time-series imagery samples. The sequence length was set as three. Thus, four sequences of satellite imagery (i.e., 1-2-3, 2-3-4, 3-4-5 and 4-5-6) can be generated for each quadrat in the experiment of 2016 and two sequences of satellite imagery (i.e., 3-4-5 and 4-5-6) can be generated for each quadrat in the experiment of 2017. In the training of XGBoost and the FCDNN, only labeled satellite imagery at time points 3, 4, 5 and 6 was used to train the model. For the three methods, since there were 200 quadrats for training and 40 quadrats for testing in each year, the number of training samples was 1200 (200×6) and the number of testing samples was 240 (40×6).

Table 2. Settings of different calculations (N1: Number of training samples; P1: Percentage of diseased samples in the train set; N2: Number of testing samples P2: Percentage of diseased samples in the test set).

No.	Methods	Input (Target)	N1	P1	N2	P2
I	GRU	2016: 1-2-3 (3), 2-3-4 (4), 3-4-5 (5), 4-5-6 (6) 2017: 3-4-5 (5), 4-5-6 (6)	1200	14.92%	240	19.58%
	XGBoost	2016: 3 (3), 4 (4), 5 (5), 6 (6); 2017: 5 (5), 6 (6)				
	FCDNN	2016: 3 (3), 4 (4), 5 (5), 6 (6); 2017: 5 (5), 6 (6)				
II	GRU	2016: 1-2-3-4 (4), 2-3-4-5 (5), 3-4-5-6 (6) 2017: 2-3-4-5 (5), 3-4-5-6 (6)	1000	17.90%	200	23.50%
	XGBoost	2016: 4 (4), 5 (5), 6 (6); 2017: 5 (5), 6 (6)				
	FCDNN	2016: 4 (4), 5 (5), 6 (6); 2017: 5 (5), 6 (6)				
III	GRU	2016: 1-2-3-4-5 (5), 2-3-4-5-6 (6) 2017: 1-2-3-4-5 (5), 2-3-4-5-6 (6)	800	22.25%	160	29.38%
	XGBoost	2016: 5 (5), 6 (6); 2017: 5 (5), 6 (6)				
	FCDNN	2016: 5 (5), 6 (6); 2017: 5 (5), 6 (6)				
IV	GRU	2016: 1-2-3-4-5-6 (6) 2017: 1-2-3-4-5-6 (6)	400	30%	80	40.00%
	XGBoost	2016: 6 (6); 2017: 6 (6)				
	FCDNN	2016: 6 (6); 2017: 6 (6)				

In calculation II, the target was SDS at the time points 4, 5 and 6 in 2016 and time points 5 and 6 in 2017. In the training of the GRU-based model, the sequence length was set as four. Thus, three sequences of

satellite imagery (i.e., 1-2-3-4, 2-3-4-5 and 3-4-5-6) can be generated for each quadrat in the experiment of 2016 and two sequences of satellite imagery (i.e., 2-3-4-5 and 3-4-5-6) can be generated for each quadrat in the experiment of 2017. The number of training samples was 1000 (200×5) and the number of testing samples was 200 (40×5).

In calculation III, the target was SDS at the time points 5 and 6 in 2016 and time points 5 and 6 in 2017. In the training of the GRU-based model, the sequence length was set as five. Thus, two sequences of satellite imagery (i.e., 1-2-3-4-5 and 2-3-4-5-6) can be generated for each quadrat in the experiment of 2016 and two sequences of satellite imagery (i.e., 1-2-3-4-5 and 2-3-4-5-6) can be generated for each quadrat in the experiment of 2017. The number of training samples was 800 (200×4) and the number of testing samples was 160 (40×4).

In calculation IV, the target was SDS at the time point 6 in 2016 and time point 6 in 2017. In the training of the GRU-based model, the sequence length was set as six. Thus, one sequence of satellite imagery (i.e., 1-2-3-4-5-6) can be generated for each quadrat in the experiment of 2016 and one sequence of satellite imagery (i.e., 1-2-3-4-5-6) can be generated for each quadrat in the experiment of 2017. The number of training samples was 400 (200×2) and the number of testing samples was 80 (40×2).

The results are shown in Table 3. In calculation I, the three methods had the same test precision. However, the test recall of GRU was 9% and 51% greater than that of XGBoost and the FCDNN, respectively. The reason why the test recall of the FCDNN is much less is that the model was overfitted and most of the samples were predicted as healthy. In calculation II, the test accuracy of GRU was 2% greater than that of XGBoost and 0.5% greater than that of the FCDNN owing to the improved test precision. It means that most of the positive predictions were accurate. In calculation III, the test accuracy of GRU was 1.8% and 5% greater than that of XGBoost and the FCDNN, respectively. In calculation IV, the test accuracy of GRU was 1.2% and 6% greater than that of XGBoost and the FCDNN, respectively. It can be observed that the training accuracy of GRU was lower in the four calculations; however, the test accuracy of GRU is the highest among the three methods in the four calculations. It proves the good generalization performance of the GRU-based method. The test precision of GRU was highest in the four calculations, which was about 80–95%. In terms of test recall, the FCDNN outperformed the other two methods in calculations II, III and IV. However, its test accuracy was lower because the proportion of positives that are correctly identified was lower.

Table 3. Comparisons among the three methods in different calculations.

No.	Methods	Training Accuracy	Test Accuracy	Test Precision	Test Recall
I	GRU	0.909	0.904	0.800	0.681
	XGBoost	0.947	0.892	0.800	0.596
	FCDNN	0.903	0.829	0.800	0.170
II	GRU	0.885	0.860	0.890	0.532
	XGBoost	0.928	0.840	0.742	0.489
	FCDNN	0.929	0.855	0.855	0.638
III	GRU	0.865	0.856	0.800	0.681
	XGBoost	0.931	0.838	0.784	0.617
	FCDNN	0.905	0.806	0.643	0.766
IV	GRU	0.820	0.825	0.950	0.594
	XGBoost	0.865	0.813	0.905	0.594
	FCDNN	0.853	0.763	0.697	0.719

The confusion matrices are shown in Figure 7. In calculation I, although GRU has the lowest number of true predictions for healthy quadrats, it has more true predictions for diseased quadrats than the other two methods. For the FCDNN, the opposite is the case, since 230 quadrats were classified as healthy and only 10 quadrats were classified as diseased. The reason is that the FCDNN is more likely to predict samples as healthy. In calculation II, GRU made more true predictions for healthy quadrats. In calculation

III, Both GRU and XGBoost classified 105 healthy quadrats correctly; however, GRU had more true predictions for diseased quadrats than XGBoost. In calculation IV, the performance of GRU was similar to that of XGBoost; the difference was only one healthy quadrat. The FCDNN did not perform well in the classification of healthy quadrats. In conclusion, the calculation results show that prediction accuracy can be improved by using a sequence-based model, i.e., GRU, with time-series imagery.

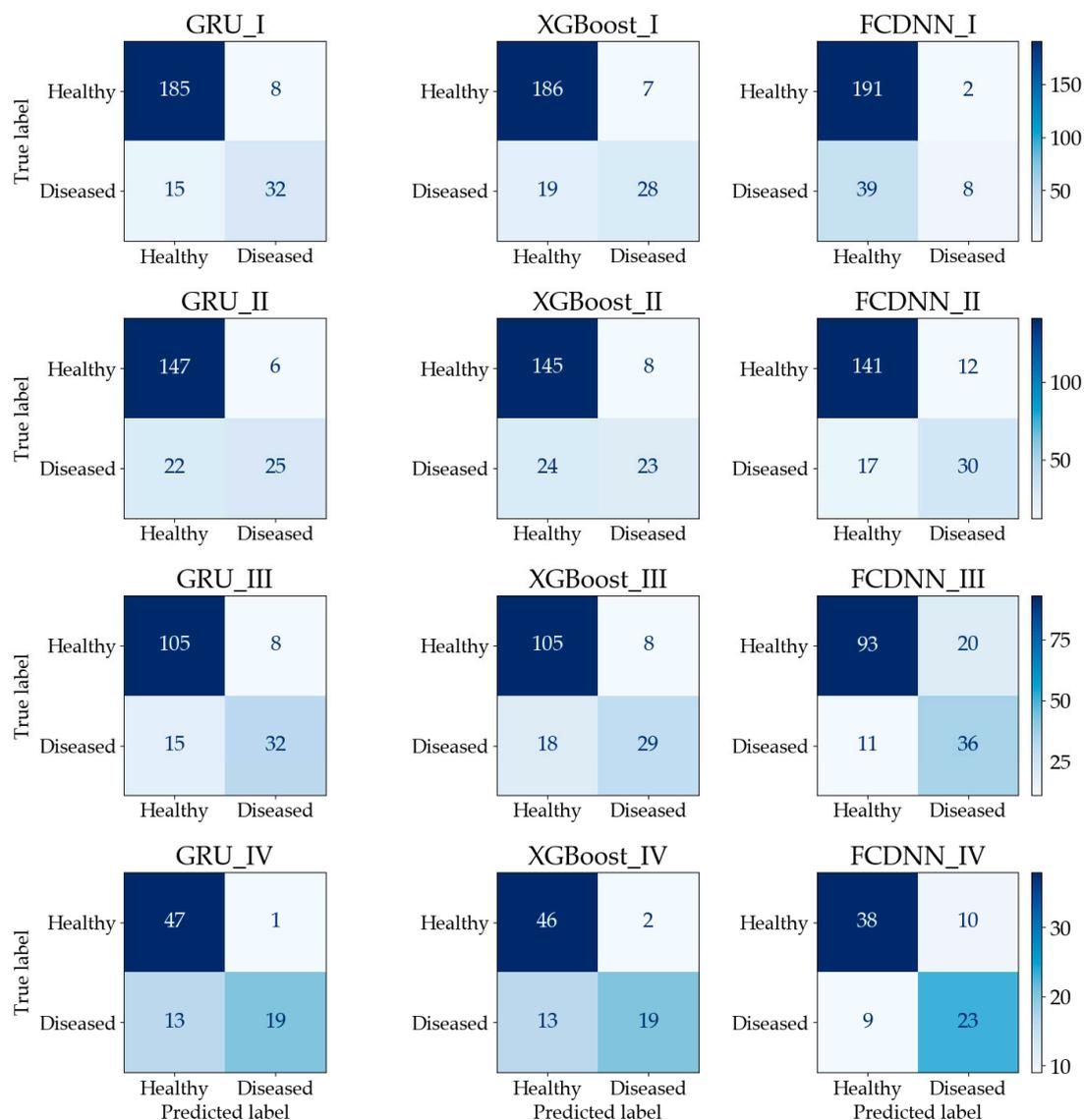


Figure 7. Confusion matrix of the testing dataset (each figure is named as “[method name]_[calculation number]”).

3.3. Data Imbalance

The number of healthy samples was much larger than that of diseased samples (Table 2). This is a typical example of a data imbalance issue. One possible consequence is that the learning models will be guided by global performance, while the minority samples may be treated as noise [49]. For example, for a dataset consisting of 90 negative examples and 10 positive samples, a classifier that predicts all samples to be positive can achieve an overall accuracy of 90%. However, the classifier does not have the ability to predict positive examples. In our case, the training processes of the models were guided by the cross-entropy loss function. To reduce the loss, models were more likely to predict the samples as healthy since healthy samples were the majority. This pattern can be observed from the

results shown in Table 3. In most of the situations, the test precision was 10–20% greater than the test recall. There are two common approaches to address the issue of data imbalance. One way is to eliminate the data imbalance by using sampling methods [50]. Over-sampling methods create more new minority classes while under-sampling methods discard some samples in the majority class [51,52]. Another way is to increase the costs for the misclassification of minority class samples [53]. Since the number of samples was limited, the second way was adopted to test model performance by assigning different weights to the minority class samples. The results are shown in Figure 8. In calculation I (Figure 8a), when the weight of the ‘disease’ class was increased from 1 to 1.2, the results did not change. When the class weight was between 1.2 and 2.8, with the increase in class weight, the recall of diseased samples increased from 0.681 to 0.830. However, the overall test accuracy and the test precision dropped by 15% and 35%, respectively. In calculation II (Figure 8b), when the weight was 1.8, the test recall achieved 0.830. After that, the model performance deteriorated with the increase in the class weight. In calculation III (Figure 8c), the best test recall was achieved when the class weight was increased to 2.6. The overall test accuracy and the test precision dropped by 20% and 34%, respectively. In calculation IV (Figure 8d), the test recall can be improved to 0.8. In summary, the increase in minority class weight can help improve the classification accuracy of diseased samples. However, the improvement was at the cost of overall accuracy and precision. Therefore, the value of the class weight should be adjusted according to the practical need. If the objective is to detect as many diseased quadrats as possible, a larger value of class weight should be used; otherwise, a smaller value should be used.

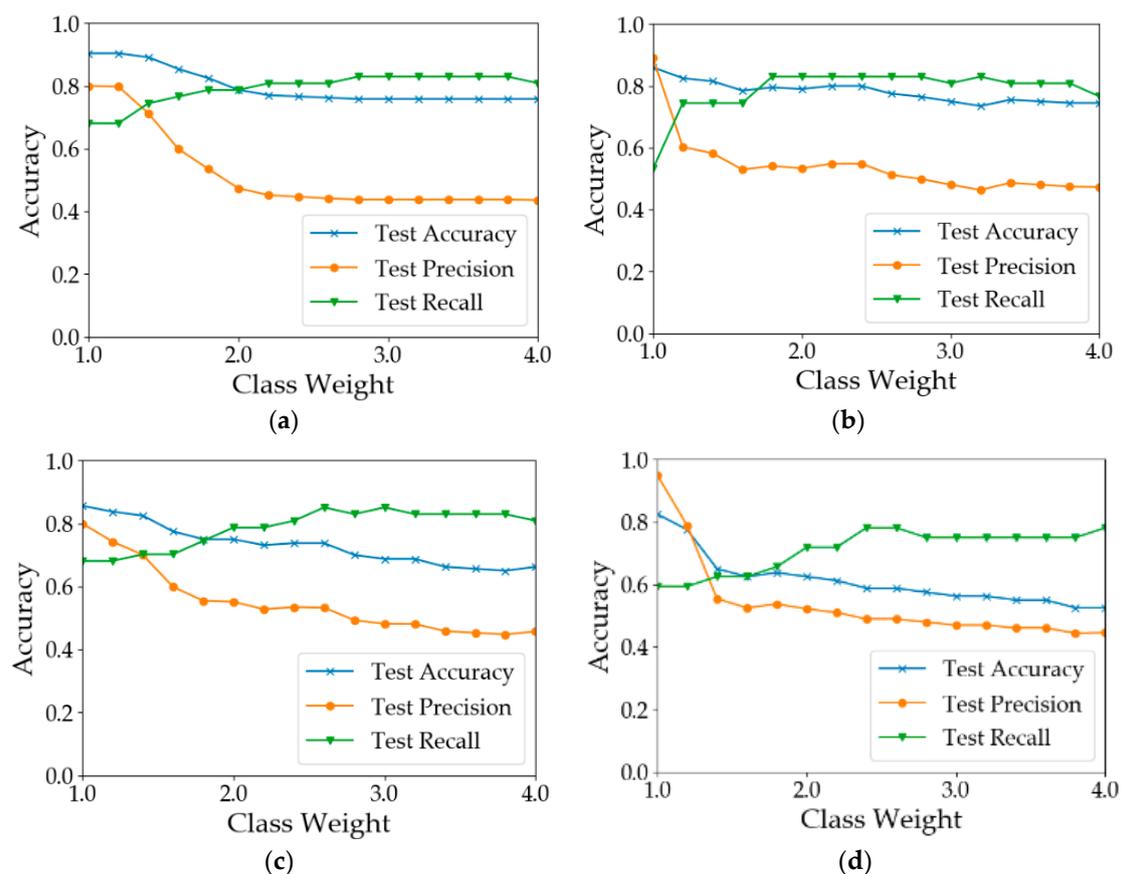


Figure 8. Test accuracy, precision and recall using different weights for the minority class. (a) Results of Calculation I (b) Results of Calculation II (c) Results of Calculation III (d) Results of Calculation IV

3.4. Forecast of the Soybean Sudden Death Syndrome (SDS)

Since the GRU-based method is a time sequence prediction model, it should forecast SDS occurrence in the future. To measure the prediction performance, calculations of four different scenarios have been conducted, as shown in Table 4. The first method was named as GRU_Current, which is the same as the proposed method in Section 3.2. The second was named as GRU_Next, which uses time-series imagery to predict SDS at the next time point. The results have been compared with the results in Table 3. Take calculation A as an example; the target was SDS at time points 3, 4, 5 and 6 in 2016 and time points 5 and 6 in 2017. The sequence length of GRU_Next was two. The sequences were 1-2, 2-3, 3-4 and 4-5 in 2016 and 3-4 and 4-5 in 2017. The sequence length of GRU_Current was three. The sequences were 1-2-3, 2-3-4, 3-4-5 and 4-5-6 in 2016 and 3-4-5 and 4-5-6 in 2017. The two methods have the same target, i.e., SDS at the time points 3, 4, 5 and 6 in 2016 and at time points 5 and 6 in 2017.

Table 4. Input imagery sequence and target dates of the two methods. The values in parenthesis indicate target dates.

No.	Input (target) of GRU_Current	Input (target) of GRU_Next
A	2016: 1-2-3 (3), 2-3-4 (4), 3-4-5 (5), 4-5-6 (6) 2017: 3-4-5 (5), 4-5-6 (6)	2016: 1-2 (3), 2-3 (4), 3-4 (5), 4-5 (6) 2017: 3-4 (5), 4-5 (6)
B	2016: 1-2-3-4 (4), 2-3-4-5 (5), 3-4-5-6 (6) 2017: 2-3-4-5 (5), 3-4-5-6 (6)	2016: 1-2-3 (4), 2-3-4 (5), 3-4-5 (6) 2017: 2-3-4 (5), 3-4-5 (6)
C	2016: 1-2-3-4-5 (5), 2-3-4-5-6 (6) 2017: 1-2-3-4-5 (5), 2-3-4-5-6 (6)	2016: 1-2-3-4 (5), 2-3-4-5 (6) 2017: 1-2-3-4 (5), 2-3-4-5 (6)
D	2016: 1-2-3-4-5-6 (6) 2017: 1-2-3-4-5-6 (6)	2016: 1-2-3-4-5 (6) 2017: 1-2-3-4-5 (6)

The results are shown in Table 5. In calculations A, B and C, the test accuracy of GRU_Current was 1–3% greater than that of GRU_Next. In terms of test precision and test recall, GRU_Current also outperformed GRU_Next. However, with the increase in the sequence length, the gap between the test accuracy of the two methods becomes smaller. In calculation D, although the train accuracies were slightly different, the test accuracy, recall and precision of the two methods were the same. It indicates that the proposed model can be used to predict the SDS at the next time point when enough historical images are available, which will bring benefits to the prediction of the future development of the plant diseases.

Table 5. Comparisons between two methods.

Methods	No.	Train Accuracy	Test Accuracy	Test Precision	Test Recall
GRU_Current	A	0.909	0.904	0.800	0.681
GRU_Next	A	0.885	0.875	0.718	0.596
GRU_Current	B	0.885	0.860	0.890	0.532
GRU_Next	B	0.871	0.830	0.760	0.404
GRU_Current	C	0.865	0.856	0.800	0.681
GRU_Next	C	0.864	0.844	0.789	0.638
GRU_Current	D	0.820	0.825	0.950	0.594
GRU_Next	D	0.830	0.825	0.950	0.594

4. Discussion

Remote sensing using satellite imagery can be a potentially powerful tool to detect plant diseases at a quadrat or field level. Our results prove that the stress-triggered changes in the pattern of light emission due to soybean SDS can be detected through high-resolution satellite imagery and the

classification accuracy of diseased and healthy quadrats can be further improved by incorporating time-series prediction. Our proposed method has manifested its ability to improve SDS detection accuracy by incorporating time-series information.

A growing interest has been observed recently in the early detection of plant diseases. Researchers have used different remote sensing tools for early detection and monitoring of plant diseases at different spatial levels, such as leaf scale [54–59], plant canopy scale [3,60], plot-scale based on aerial imagery [61,62] and field-wide-scale based on satellite imagery [29,63–65]. Although many studies obtained remotely sensed data at different time points, none of them incorporated time-series information in the analytical models.

Instead of only using individual static imagery, we proposed a model that treats the satellite imagery, captured at different time points, as a time series for the detection of SDS. We compared the SDS prediction results from our proposed GRU-based method with XGBoost and the FCDNN, both non-sequence-based methods. Although the test accuracies of all three methods were above 76%, accuracy improved by up to 7% after incorporating time-series prediction. This substantial improvement in accuracy reveals that the GRU-based method uses the characteristic information of spectral bands, ground-based crop rotation and time series in an optimal way. The main advantage of the GRU-based method goes back to its workability of learning from history data. In the learning process, GRU can determine the influence of images at different time points on the decision-making of the current status through the reset gate and the update gate. Information from past images can help the model to eliminate the effect of some noise, such as weather conditions.

In all calculation scenarios, the GRU-based method outperformed XGBoost and the FCDNN in SDS prediction accuracy. Although XGBoost and the FCDNN are very powerful methods, they do not incorporate time-based information. In our data, we found that reflectance values in the RGB spectrum were lower for healthy quadrats than diseased quadrats, while this was the opposite in the NIR spectrum. Moreover, this pattern became clearer near the end of the cropping season, which indicates that incorporating time-based information for satellite images can add information to the data analysis.

Besides accuracy, the GRU-based method also achieved greater precision (80–95%) for SDS detection in all calculation scenarios, as compared to the XGBoost and FCDNN methods. This means that soybean quadrats predicted either as diseased or healthy were assigned correctly 80% to 95% of the time. However, in terms of recall, the FCDNN performed better than the GRU-based method and XGBoost in three out of four calculation scenarios. This means that the FCDNN predicted diseased soybean quadrats more accurately than the other two methods. However, its test accuracy was lower because this method was not correctly predicting the healthy soybean quadrats. The reason is that FCDNN was more likely to predict a sample as diseased, so the precision was sacrificed.

In the end, we made predictions of SDS at future time points (GRU_Next) based on previous time-based imagery and we compared these results with the predictions made at time points included in the training (GRU_Current). SDS prediction accuracy, precision and recall were greater in GRU_Current models in small sequence scenarios. It is possible that this improved precision is because the models in the GRU_Current scenarios can extract information from the imagery collected at the current time point. On the other hand, in GRU_Next scenarios, we were predicting SDS using the history data only. Moreover, in the last sequence scenario, accuracy, recall and precision of both GRU_Next and GRU_Current models became equal, which indicates that the GRU-based method can predict SDS in soybean quadrats with high accuracy when enough historical images are available.

A recent study [29] conducted at the same site detected SDS through satellite images with above 75% accuracy using the random forest algorithm. Contrary to our study, they used static images for data analysis and predicted SDS on the 30% test subset from the same satellite images. In comparison, we proposed a method that can predict SDS with high accuracy at future time points, based on time-series satellite imagery.

5. Limitations and Future Work

It should be noted that this proposed plant-disease recognition using a sequence-based model with remote sensing is subject to a few limitations which suggest future research directions.

In the data collection, the number of images in our case study was limited because the PlanetScope satellite can only take pictures of the locations every one or two weeks. More frequent observations of SDS can help the model capture subtle changes along time. Some studies installed a camera system around the canopies or plots so that they can monitor the real-time status of the plants. However, since it requires a high intensity of cameras, the scalability of this method is limited. Another issue of satellite imagery is that the collection frequency of the imagery may not correspond to the development of diseases. As shown in the previous section, SDS development in June and July was slow, while it was much faster at the end of August. So, ideally, it would have been better to collect images more frequently in August. Besides, the locations of the quadrats of 2016 and 2017 were different. Therefore, the time sequences of 2016 and 2017 were constructed separately. In the future, the experimental fields should be evaluated consistently so that longer time sequences can be used to further improve the prediction accuracy.

In terms of data preprocessing, this paper only used the mean values and variance values of the pixels due to the resolution limitations. The importance of different pixels may be different due to disease intensities. One alternative way is to construct a small-scale CNN to obtain features of each image and then feed the extracted features to the GRU. Additionally, multi-band imagery can also be used to improve the accuracy.

In some calculations, the percentage of diseased samples was very low, meaning that the data were not balanced. In this paper, this issue was addressed by assigning weights to the diseased samples in the calculation of loss of function. There are also some other methods to address data imbalance problem. For example, the oversampling method can generate new diseased samples by recombining the pixels of the diseased samples. In the future, we will compare the effectiveness of different methods.

Lastly, in this study, SDS is labeled as diseased (denoted as “1”) and not diseased (denoted as “0”), using 5% as a threshold value which constitutes a classification problem. One concern is the influence of the threshold value on the model performance. For example, in our case, quadrats with 5.1% and 99.1% incidence were categorized as diseased samples, which may influence the prediction precision of the model. Therefore, the classification accuracy of the model will be tested using different threshold values in the future. We can also test other time-series models, e.g., LSTM. These are reserved for future research.

6. Conclusions

The development of sensing techniques has brought significant improvements to plant-disease detection, treatment and management. However, there is a lack of research on detecting SDS using pixel-level satellite imagery at the quadrat level. The major challenge is how to use limited information, i.e., only a few pixels for each quadrat, to improve SDS prediction accuracy. Some traditional methods for the analysis of near-sensing data are based on CNNs, which can efficiently extract the most important features from thousands of RGB values and other information. In contrast, for the analysis of low-pixel-level satellite imagery, additional information is required.

In this paper, a GRU-based model is proposed to predict SDS by incorporating temporality into the satellite imagery. Instead of using individual static imagery, time-series imagery is used to train the model. Different test case scenarios have been created for comparisons between the proposed method and other non-sequence-based methods. The results show that compared to XGBoost and the FCDNN, the GRU-based method can improve overall prediction accuracy by 7%. In addition, the proposed method can potentially be adapted to predict future development of SDS.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/12/21/3621/s1>, The satellite data collected at Marsden Farm in 2016 and 2017 are available online at https://github.com/LnB11855/remote_sensing_paper. This repository also contains the Python code of data processing and model training.

Author Contributions: Conceptualization, L.B., Y.K., and G.H.; methodology, L.B. and G.H.; data collection, M.M.R.; writing—original draft preparation, L.B.; writing—review and editing, G.H., M.M.R., Y.K., L.L. and D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the soybean checkoff through the United Soybean Board and the USDA-NIFA Crop Protection and Pest Management Program, grant number 2019-02989.

Acknowledgments: We thank Sharon Eggenburger from the Department of Plant Pathology and Microbiology for her help in data collection and Matt Liebman from the Department of Agronomy for allowing us to use the ongoing soybean field experiment located at the Iowa State University Marsden Farm.

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

References

- Allen, T.W.; Bradley, C.A.; Sisson, A.J.; Byamukama, E.; Chilvers, M.I.; Coker, C.M.; Collins, A.A.; Damicone, J.P.; Dorrance, A.E.; Dufault, N.S. Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. *Plant Health Prog.* **2017**, *18*, 19–27. [[CrossRef](#)]
- Crop Protection Network. Estimates of Corn and Soybean Yield Losses Due to Disease: An Online Tool. 2020. Available online: <https://loss.cropprotectionnetwork.org/Doi.org/10.31274/cpn-20191121-0> (accessed on 15 July 2020).
- Herrmann, I.; Vosberg, S.K.; Ravindran, P.; Singh, A.; Chang, H.-X.; Chilvers, M.I.; Conley, S.P.; Townsend, P.A. Leaf and canopy level detection of *Fusarium virguliforme* (sudden death syndrome) in soybean. *Remote Sens.* **2018**, *10*, 426. [[CrossRef](#)]
- Barton, C.V.M. Advances in remote sensing of plant stress. *Plant Soil* **2012**, *354*, 41–44. [[CrossRef](#)]
- Mahlein, A.-K. Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* **2016**, *100*, 241–251. [[CrossRef](#)] [[PubMed](#)]
- Nilsson, H. Remote sensing and image analysis in plant pathology. *Annu. Rev. Phytopathol.* **1995**, *33*, 489–528. [[CrossRef](#)] [[PubMed](#)]
- Zwiggelaar, R. A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. *Crop Prot.* **1998**, *17*, 189–206. [[CrossRef](#)]
- Jacquemoud, S.; Ustin, S.L. Leaf optical properties: A state of the art. In Proceedings of the 8th International Symposium of Physical Measurements & Signatures in Remote Sensing, Aussois, France, 8–12 January 2001; pp. 223–332.
- Chaerle, L.; Van Der Straeten, D. Imaging techniques and the early detection of plant stress. *Trends Plant Sci.* **2000**, *5*, 495–501. [[CrossRef](#)]
- Mastrodimos, N.; Lentzou, D.; Templalexis, C.; Tsitsigiannis, D.; Xanthopoulos, G. Development of thermography methodology for early diagnosis of fungal infection in table grapes: The case of *Aspergillus carbonarius*. *Comput. Electron. Agric.* **2019**, *165*, 104972. [[CrossRef](#)]
- Wang, Y.; Zia-Khan, S.; Owusu-Adu, S.; Miedaner, T.; Müller, J. Early Detection of *Zymoseptoria tritici* in Winter Wheat by Infrared Thermography. *Agriculture* **2019**, *9*, 139. [[CrossRef](#)]
- Bürbling, K.; Hunsche, M.; Noga, G. Use of blue–green and chlorophyll fluorescence measurements for differentiation between nitrogen deficiency and pathogen infection in winter wheat. *J. Plant Physiol* **2011**, *168*, 1641–1648. [[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](#)]
- Pérez-Bueno, M.L.; Pineda, M.; Barón Ayala, M. Phenotyping plant responses to biotic stress by chlorophyll fluorescence imaging. *Front. Plant Sci.* **2019**, *10*, 1135. [[CrossRef](#)]
- Nagasubramanian, K.; Jones, S.; Sarkar, S.; Singh, A.K.; Singh, A.; Ganapathysubramanian, B. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant Methods* **2018**, *14*, 86. [[CrossRef](#)]

16. 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](#)] [[PubMed](#)]
17. Förster, A.; Behley, J.; Behmann, J.; Roscher, R. Hyperspectral plant disease forecasting using generative adversarial networks. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 1793–1796.
18. Durmuş, H.; Güneş, E.O.; Kırıcı, M. Disease detection on the leaves of the tomato plants by using deep learning. In Proceedings of the 2017 6th International Conference on Agro-Geoinformatics, Fairfax, VA, USA, 7–10 August 2017; pp. 1–5.
19. 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](#)]
20. Liu, W.; Cao, X.; Fan, J.; Wang, Z.; Yan, Z.; Luo, Y.; West, J.S.; Xu, X.; Zhou, Y. Detecting wheat powdery mildew and predicting grain yield using unmanned aerial photography. *Plant Dis.* **2018**, *102*, 1981–1988. [[CrossRef](#)]
21. Bajwa, S.; Rupe, J.; Mason, J. Soybean disease monitoring with leaf reflectance. *Remote Sens.* **2017**, *9*, 127. [[CrossRef](#)]
22. Hatton, N.M.; Menke, E.; Sharda, A.; van der Merwe, D.; Schapaugh, W., Jr. Assessment of sudden death syndrome in soybean through multispectral broadband remote sensing aboard small unmanned aerial systems. *Comput. Electron. Agric.* **2019**, *167*, 105094. [[CrossRef](#)]
23. Hatton, N.M. Use of Small Unmanned Aerial System for Validation of Sudden Death Syndrome in Soybean through Multispectral and Thermal Remote Sensing. Ph.D. Thesis, Department of Biological & Agricultural Engineering, College Station, TX, USA, 2018.
24. Yuan, L.; Pu, R.; Zhang, J.; Wang, J.; Yang, H. Using high spatial resolution satellite imagery for mapping powdery mildew at a regional scale. *Precis. Agric.* **2016**, *17*, 332–348. [[CrossRef](#)]
25. Satapathy, R.R. Remote sensing in plant disease management. *J. Pharmacogn. Phytochem.* **2020**, *9*, 1813–1820.
26. Mcgibney, A.; Klepal, M.; Pesch, D. Agent-Based Optimization for Large Scale WLAN Design. *IEEE Trans. Evol.* **2011**, *15*, 470–486.
27. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [[CrossRef](#)]
28. 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](#)]
29. Raza, M.M.; Harding, C.; Liebman, M.; Leandro, L.F. Exploring the Potential of High-Resolution Satellite Imagery for the Detection of Soybean Sudden Death Syndrome. *Remote Sens.* **2020**, *12*, 1213. [[CrossRef](#)]
30. Samajpati, B.J.; Degadwala, S.D. Hybrid approach for apple fruit diseases detection and classification using random forest classifier. In Proceedings of the 2016 International Conference on Communication and Signal Processing (ICCSP), Madras, India, 6–8 April 2016; pp. 1015–1019.
31. Chaudhary, A.; Kolhe, S.; Kamal, R. An improved random forest classifier for multi-class classification. *Inf. Process. Agric.* **2016**, *3*, 215–222. [[CrossRef](#)]
32. Probst, P.; Wright, M.N.; Boulesteix, A.L. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2019**, *9*, e1301. [[CrossRef](#)]
33. Zhang, Y.; Jiang, D.; Wang, J. A recurrent neural network for solving Sylvester equation with time-varying coefficients. *IEEE Trans. Neural Netw. Learn. Syst.* **2002**, *13*, 1053–1063. [[CrossRef](#)]
34. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
35. Sundermeyer, M.; Schlüter, R.; Ney, H. LSTM neural networks for language modeling. In Proceedings of the Thirteenth annual conference of the international speech communication association, Portland, OR, USA, 9–13 September 2012.
36. Turkoglu, M.; Hanbay, D.; Sengur, A. Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests. *J. Ambient Intell. Humaniz. Comput.* **2019**, 1–11. [[CrossRef](#)]
37. Namin, S.T.; Esmailzadeh, M.; Najafi, M.; Brown, T.B.; Borevitz, J.O. Deep phenotyping: Deep learning for temporal phenotype/genotype classification. *Plant Methods* **2018**, *14*, 66. [[CrossRef](#)]
38. Guillen-Navarro, M.A.; Martinez-Espana, R.; aBelen Auiso aJose, A.B.-C.; Moreno, L. An LSTM Deep Learning Scheme for Prediction of Low Temperatures in Agriculture. In Proceedings of the 15th International Conference on Intelligent Environments 2019, Rabat, Morocco, 24–27 June 2019; Volume 26, pp. 130–138.

39. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.
40. Jin, X.; Jie, L.; Wang, S.; Qi, H.J.; Li, S.W. Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field. *Remote Sens.* **2018**, *10*, 395. [[CrossRef](#)]
41. Liebman, M.; Gibson, L.R.; Sundberg, D.N.; Heggenstaller, A.H.; Westerman, P.R.; Chase, C.A.; Hartzler, R.G.; Menalled, F.D.; Davis, A.S.; Dixon, P.M. Agronomic and economic performance characteristics of conventional and low-external-input cropping systems in the central Corn Belt. *Agron. J.* **2008**, *100*, 600–610. [[CrossRef](#)]
42. Davis, A.S.; Hill, J.D.; Chase, C.A.; Johanns, A.M.; Liebman, M. Increasing cropping system diversity balances productivity, profitability and environmental health. *PLoS ONE* **2012**, *7*, e47149. [[CrossRef](#)]
43. Gómez, R.; Liebman, M.; Sundberg, D.N.; Chase, C.A. Comparison of crop management strategies involving crop genotype and weed management practices in conventional and more diverse cropping systems. *Renew. Agric. Food Syst.* **2013**, *28*, 220–233. [[CrossRef](#)]
44. Leandro, L.F.; Robertson, A.E.; Mueller, D.S.; Yang, X.-B. Climatic and environmental trends observed during epidemic and non-epidemic years of soybean sudden death syndrome in Iowa. *Plant Health Prog.* **2013**, *14*, 18. [[CrossRef](#)]
45. Lukoševičius, M.; Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **2009**, *3*, 127–149. [[CrossRef](#)]
46. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
47. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13 August 2016; pp. 785–794.
48. Safavian, S.R.; Landgrebe, D. A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 660–674. [[CrossRef](#)]
49. Haixiang, G.; Yijing, L.; Shang, J.; Mingyun, G.; Yuanyue, H.; Bing, G. Learning from class-imbalanced data: Review of methods and applications. *Expert Syst. Appl.* **2017**, *73*, 220–239. [[CrossRef](#)]
50. Thabtah, F.; Hammoud, S.; Kamalov, F.; Gonsalves, A. Data imbalance in classification: Experimental evaluation. *Inf. Sci.* **2020**, *513*, 429–441. [[CrossRef](#)]
51. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *Int. J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
52. Kotsiantis, S.; Kanellopoulos, D.; Pintelas, P. Handling imbalanced datasets: A review. *GESTS Int. Trans. Comput. Sci. Eng.* **2006**, *30*, 25–36.
53. Japkowicz, N.; Stephen, S. The class imbalance problem: A systematic study. *Intell. Data Anal* **2002**, *6*, 429–449. [[CrossRef](#)]
54. Rumpf, T.; Mahlein, A.K.; Steiner, U.; Oerke, E.-C.; Dehne, H.W.; Plümer, L. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Comput. Electron. Agric.* **2010**, *74*, 91–99. [[CrossRef](#)]
55. Al-Hiary, H.; Bani-Ahmad, S.; Reyalat, M.; Braik, M.; ALRahamneh, Z. Fast and accurate detection and classification of plant diseases. *Mach. Learn.* **2011**, *17*, 31–38. [[CrossRef](#)]
56. Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W.B. Early detection of Fusarium infection in wheat using hyper-spectral imaging. *Comput. Electron. Agric.* **2011**, *75*, 304–312. [[CrossRef](#)]
57. Rahman, H.U.; Ch, N.J.; Manzoor, S.; Najeeb, F.; Siddique, M.Y.; Khan, R.A. A comparative analysis of machine learning approaches for plant disease identification. *Adv. Life Sci.* **2017**, *4*, 120–126.
58. Lu, J.; Ehsani, R.; Shi, Y.; Castro, A.I.; Wang, S. Detection of multi-tomato leaf diseases (late blight, target and bacterial spots) in different stages by using a spectral-based sensor. *Sci. Rep.* **2018**, *8*, 1–11. [[CrossRef](#)]
59. Maniyath, S.R.; Vinod, P.; Niveditha, M.; Pooja, R.; Shashank, N.; Hebbar, R. Plant Disease Detection Using Machine Learning. In Proceedings of the 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, India, 25–28 April 2018; pp. 41–45.
60. Bravo, C.; Moshou, D.; West, J.; McCartney, A.; Ramon, H. Early disease detection in wheat fields using spectral reflectance. *Biosyst. Eng.* **2003**, *84*, 137–145. [[CrossRef](#)]

61. Hatton, N.; Sharda, A.; Schapaugh, W.; Van der Merwe, D. Remote thermal infrared imaging for rapid screening of sudden death syndrome in soybean. In Proceedings of the 2018 ASABE Annual International Meeting, Detroit, MI, USA, 29 July–1 August 2018; p. 1.
62. Pham, H.; Lim, Y.; Gardi, A.; Sabatini, R.; Pang, E. A novel bistatic lidar system for early-detection of plant diseases from unmanned aircraft. In Proceedings of the 31th Congress of the International Council of the Aeronautical Sciences (ICAS 2018), Belo Horizonte, Brazil, 9–14 September 2018.
63. Yuan, L.; Zhang, J.; Shi, Y.; Nie, C.; Wei, L.; Wang, J. Damage mapping of powdery mildew in winter wheat with high-resolution satellite image. *Remote Sens.* **2014**, *6*, 3611–3623. [[CrossRef](#)]
64. Zheng, Q.; Huang, W.; Cui, X.; Shi, Y.; Liu, L. New spectral index for detecting wheat yellow rust using Sentinel-2 multispectral imagery. *Sensors* **2018**, *18*, 868. [[CrossRef](#)]
65. Yang, S.; Li, X.; Chen, C.; Kyveryga, P.; Yang, X.B. Assessing field-specific risk of soybean sudden death syndrome using satellite imagery in iowa. *Phytopathology* **2016**, *106*, 842–853. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).