Journal of Food Process Engineering

WILEY

Journal of Food Process Engineering

ORIGINAL ARTICLE OPEN ACCESS

Diagnosing Fungal Infection in Wheat Kernels by Integrating Spectroscopic Technology and Digital Color Imaging System: Artificial Neural Network, Principal Component Analysis and Correlation Feature Selection Techniques

Saman Zohrabi[1,2] | Seyed Sadegh Seiiedlou[1] | Iman Golpour[3] | Mark Lefsrud[2] | Raquel P. F. Guiné[4] | Barbara Sturm[5,6]

[1]Department of Biosystems Engineering, University of Tabriz, Tabriz, Iran | [2]Department of Bioresource Engineering, Macdonald Campus, McGill University, Sainte-Anne-de-Bellevue, Quebec, Canada | [3]Department of Energy Engineering, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain | [4]CERNAS Research Centre, Polytechnic Institute of Viseu, Viseu, Portugal | [5]Leibniz Institute for Agricultural Engineering and Bioeconomy, Potsdam, Germany | [6]Albrecht Daniel Thaer-Institute of Agricultural and Horticultural Sciences, Humboldt Universität zu Berlin, Berlin, Germany

Correspondence: Saman Zohrabi (zohrabi.saman@gmail.com) | Barbara Sturm (bsturm@atb-potsdam.de)

Received: 14 August 2024 | **Revised:** 1 October 2024 | **Accepted:** 20 October 2024

Funding: The authors received no specific funding for this work.

Keywords: artificial neural network (ANN) | correlation feature selection method (CFS) | fungi infection | principal component analysis (PCA) | spectroscopy technique | wheat

ABSTRACT

Contamination of cereal grain, especially wheat, with fungal infections can cause significant economic impacts and it endangers the health of humans and livestock. This study aims to appraise the UV/VIS–NIR and digital color (RGB) imaging systems and spectroscopic methodology to detect wheat kernels infected by fungi such as *Penicillium expansum* and *Fusarium graminearum*. NIR spectra of 190–1100 nm at 10 nm intervals, visible color reflectance images and non-visible reflectance images of wheat kernels in the ultraviolet and near-infrared ranges were applied to develop the multi-layer perceptron (MLP) artificial neural network model. The optimum wavelengths were selected by application of the principal component analysis (PCA) after preprocessing the raw spectra. A confusion matrix was used in the correlation feature selection method (CFS) for the decision tree classifier of selected features. The results showed that the four UV wavelengths of 310, 330, 400, and 410 nm were the best wavelengths using PCA to distinguish healthy and unhealthy wheat kernels. Considering the intensity of the wavelengths as the neural network inputs, samples were classified into healthy and unhealthy categories with an accuracy of 90.9%. Also, 18 features of color images in RGB, LAB, HSV, HSI, YCbCr, and YIQ spaces provided the highest average accuracy of 44.4% in classifying healthy and infected wheat kernels by using a CCD Proline camera in the ultraviolet range. In contrast, other cameras in the visible and invisible range showed low accuracy. Furthermore, the best classification accuracy of the healthy and infected samples by the use of the CFS method was obtained at 88.1%. Based on the findings, spectroscopic methodology proved to be highly effective for detecting, classifying and automatic cleaning of various agricultural seeds, with a particular emphasis on wheat kernals.

1 | Introduction

Cereal grains are the most important agricultural product in the world in terms of production and cultivated area. They possess high economic value and great importance in providing staple food in human and animal life due to their nutritional composition such as vitamins, carbohydrates, fats, proteins, and minerals (Liu et al. 2022). However, the quality and safety of grains, especially wheat for consumers, is of great importance following the economic growth and social progress of countries worldwide. Wheat (*Triticum aestivum*), as a major global agricultural crop, is applied as raw material in food processing industries for making bread, cakes, cookies, pastries, crackers, pasta products, and animal feed to name but a few. The quality of these end-products is highly influenced by the quality of wheat (Dal-Pastro et al. 2016; Qin 2010; Schmidt et al. 2016). According to the FAO statistics, 10.1 million tons and 780 million tons of wheat were produced in Iran and the world in 2021, respectively (FAO 2021).

There are over 300 types of mycotoxins produced by *Fusarium*, *Aspergillus*, and *Penicillium* fungus species, which are growing in a wide variety of agricultural products during both the pre-harvest and post-harvest periods (Hedayati and Mohammadpour 2005). Products infected by these fungi and contaminated by the mycotoxins as their digestive products may pose potent hazards such as acute liver damage, Cirrhosis of the liver, induction of tumors and teratogenic effects in humans or animals because of their mycotoxins-producing character (Hedayati and Mohammadpour 2005). Fungal invasion and toxin production during post-harvest stages, including storage, processing, transportation, and sale, can occur due to factors such as moisture, temperature, physical damage, poor hygiene, improper storage, insect infestation, crop variety susceptibility, and delayed harvesting, all promoting favorable conditions for fungal growth and mycotoxin contamination (Wagacha and Muthomi 2008). Thus, efficiently detecting, identifying and separating the samples that are contaminated with fungi is of great importance in order to reduce the risk of mycotoxins entering the food chain. Traditionally, fungal contamination in agricultural products is determined using microbiological methods in a laboratory setting, which includes microbial incubation, the total mildew count, preparation slide culture methods, fungal enumerating using plate-counting or direct plating techniques, isolating in appropriate media and identifying the genus and species level by morphological characterization, including macroscopic characteristics (color, size, and colony appearance) and microscopic characteristics (conidia, conidiophore, and conidial heads) (Hedayati and Mohammadpour 2005). These methods for determining fungal-infected wheat kernels may deliver accurate results in laboratories. However, they are expensive and time-consuming, and most of them require skilled personnel and a well-equipped laboratory. They are also destructive to the test samples, making them impossible for large-scale nondestructive screening detection or integration in an online sorting and production system.

Therefore, in the last decades, the demand for developing a rapid and non-destructive method for sensing fungal-infected wheat kernels that is suitable for real-time and online detection has received significant attention. Among modern engineering methods, optical-based methods have been reported to demonstrate outstanding potential for online applications (Beale, Hagan, and Demuth 2012). The application of spectral characteristics is apparent from given differences between the chemical composition of defective and healthy kernels, and so a hyperspectral imaging (HSI) system would be beneficial to distinguish healthy seeds from fungal-damaged seeds (Singh et al. 2012). The new techniques based on morphological and color characteristics through imaging within the visible spectrum, focusing on features such as shape and size (Alibeyglu, Ghaffari, and Alipasandi 2013; Alipasandi, Ghaffari, and Alibeyglu 2013; Majumdar and Jayas 2000; Visen et al. 2004; Zohrabi, Seiiedlou, Safari, et al. 2015), HSI and ultraviolet, visible, and near-infrared spectroscopy (UV/VISS and NIRS) (Qin 2010; Zhao et al. 2020; Zohrabi, Seiiedlou, Hampanejad, et al. 2015) have provided compelling and promising results for determining the quality of wheat and fungi contamination. The HSI method was employed to detect the defects and contaminants in different agricultural products, for instance, cereal grains (Jia et al. 2020), milled grain (Femenias et al. 2021; Parrag et al. 2020; Zhao et al. 2020), maize seeds (Yang et al. 2018), tomatoes (Brdar et al. 2021) and banana (Chu et al. 2023). Ma, Ji, and Lee (2016) employed spectra with textural features extracted from the HSI technique to recognize citrus yellow dragon disease, so that the classification accuracy of three samples obtained 100%, 93.3%, and 92.9%, respectively. Sha, Hu, and Weng (2023) applied the color and HSI system to detect the infection of black root mold in apple fruits. They developed the classification models using a support vector (SV) machine, K-nearest neighbor, and random forest (RF) and concluded that random forest had the best optimal results based on the 100% accuracy for training data and 96% accuracy for prediction set. Almoujahed et al. (2024) studied the severity of Fusarium head blight through visible near-infrared (vis–NIR) spectroscopy, with wavelengths ranging from 400 to 1700 nm, and mid-infrared (MIR) spectroscopy, spanning wavenumbers from 4000 to 650 cm^{-1}, to predict Fusarium head blight infection in wheat kernels and flour. They employed two machine learning algorithms: random forest and linear discriminant analysis. Both models demonstrated higher test accuracy for flour samples—96.6% and 100%, respectively—compared to 93.1% accuracy for kernels using MIR spectroscopy. Aflatoxin and fumonisin infection identification and spectral classification of corn kernels were accomplished by Chavez et al. (2022). Corn kernels with infection of aflatoxin and fumonisin based on the 864 corn kernels extracted from nine bulk samples were separately imaged for the reflectance parameter in the ultraviolet/visible/near-infrared spectrum (304–1086 nm), then measured for aflatoxin and fumonisin by enzyme-linked immunosorbent assay. For this study, the most significant models for aflatoxin and fumonisin classification, with accuracies of 83% and 86%, were selected. Wang et al. (2005) used color digital images to determine the vitreousness of durum wheat using an artificial neural network (ANN). By applying NIR reflectance spectroscopy in wavelengths from 500 to 1700 nm for corn kernels that were infected with different conventional fungi, it was found that two NIR reflectance spectral bands centered at 715 nm and 965 nm could correctly identify 98.1% of asymptomatic kernels (Pearson and Wicklow 2006). Mustafa et al. (2024) significantly advanced the classification of Fusarium head blight in wheat by achieving a nine-level classification accuracy of over 80% across various spatiotemporal conditions. Their approach involved fusing

wavelet features and texture features derived from near-infrared HSI and non-imaging data. In another study, Xin et al. (2023) utilized the spectra and color image (RGB) features from the HSI technique to recognize the degree of infection of Botrytis disease in apple fruit. Identification models of the fungal infection degree were evaluated by input-preprocessed CNN (IPCNN), K-nearest neighbor, random forest, and convolutional neural network (CNN). The classification results showed that the fusion of efficient spectra and textural features is much better than those by application of a single feature and also IPCNN had the best identification accuracy. Peiris, Pumphrey, and Dowell (2009) examined NIR absorbance characteristics of various concentrations of *deoxynivalenol* (DON) as well as sound and *Fusarium*-damaged single wheat kernels. Wang et al. (2004) classified healthy and fungal-damaged soybean seeds and differentiated among various types of fungal damage using wavelengths between 400 and 1700 nm.

All of the above research works affirmed the detection possibility of fungal contaminations for different agricultural products by the use of the HSI method. However, there have been few reports about detecting wheat kernels infected by *Penicillium expansum* and *F. graminearum* through spectral data and analysis of ANNs based on multi-layer perceptron (MLP) algorithm, principal component analysis (PCA) and correlation feature selection (CFS) methods. Integrating these approaches offers several advantages, including enhanced accuracy, sensitivity, and specificity in fungal detection compared to traditional methods.

Hence, the main purposes of this study were: (a) To give a comprehensive investigation in the application of UV/VIS–NIR imaging and spectroscopy imaging (HSI) techniques in rapid and improved non-destructive detection of fungal contamination in wheat kernels with the help of ANNs. (b) To identify a reduced set of wavelengths to be used in the future development of a low-cost imaging system using PCA and CFS methodologies. These techniques are described in terms of their working principles, features and application advantages and disadvantages in detecting fungal contamination. In addition, perspectives on their future trends and challenges are discussed.

2 | Materials and Methods

2.1 | Sample Preparation

Iranian Alvand and Sardari wheat varieties with 15% moisture content (wet basis) were used as the samples of this study. The moisture content of wheat samples was determined by applying a standard oven method (ASAE 1988). Healthy wheat samples (approximately 50 kg of bulk wheat) were provided from the Agriculture Research Center of Miandoab, West Azerbaijan, Iran. The species of storage pure fungi, namely *P. expansum* and *F. graminearum*, were separately obtained from the Department of Plant Protection of Tabriz University. The obtained fungi were grown on a nutrient-rich medium and subsequently incubated at a temperature of 25°C for 1 week. The spore suspensions prepared from *Fusarium* and *Penicillium* were placed in plastic spray bottles. 1 kg samples of each variety of wheat kernels were sterilized by soaking in 1% sodium hypochlorite for 2 min. The wheat kernels were thoroughly rinsed in sterilized water and placed on a paper towel for 2 h. Approximately 100 g of moistened wheat was placed in a plastic tub for each fungus to induce contamination. Wheat samples were inoculated with *Fusarium* and *Penicillium* spore suspensions in a Laminar flow hood and covered with a loose plastic bag at 30°C for 2 weeks. To ensure consistency in moisture content, the same process was applied to the healthy samples, with the only difference being the omission of the inoculation step. These sound samples were also kept under the same incubation conditions, ensuring that all samples, both healthy and contaminated, had consistent moisture levels throughout the experiment.

A total of 300 wheat kernels were chosen for imaging with 50 samples, each representing healthy wheat, wheat infected by *F. graminearum* and wheat infected by *P. expansum* for both Sardari and Alvand varieties. The remaining inoculated wheat samples were processed into flour for spectroscopic analysis, comprising 14 samples of healthy flour, 14 samples of flour contaminated with *Fusarium* and 14 samples contaminated with *Penicillium* for both Alvand and Sardari varieties.

2.2 | Optical Techniques

2.2.1 | UV/VIS–NIR Imaging System

A schematic view of the imaging system is presented in Figure 1. Imaging was carried out in the range of ultraviolet, visible, and near-infrared wavelengths from 100 to 1780 nm. A 20 W fluorescent lamp was used for the ultraviolet range and a 50 W fluorescent lamp and a halogen lamp were used for visible and infrared ranges, respectively. The lamps were mounted at 45° angles and 0.5 m from the imaging area. To acquire the image in the range of UV, visible and infrared, three cameras with the following specifications were used: CCD Proline camera (PR-56551/3"SONY Super CCD, Min. Illumination 0.01 Lux, a zoom lens of 3.5–8 mm focal length), AGNI camera (AG-912XN, 1/3" Sony Super HAD CCD II, Min. Illumination 0.01 Lux a zoom lens of 3.5–8 mm Vari-focal) and LT Lutron Digital Instrument. A sample of images of fungi-infected wheat kernels in the ultraviolet range using a CCD Proline camera is presented in Figure 2.

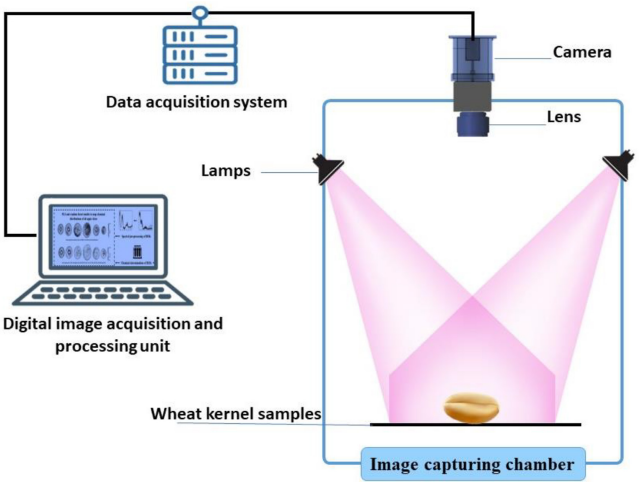

FIGURE 1 | Schematic view of the digital color imaging system.

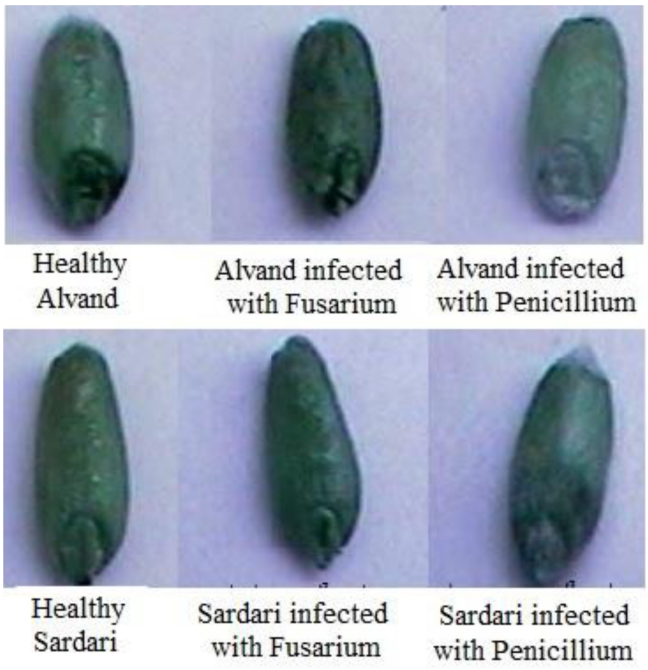

FIGURE 2 | Images of fungi infected wheat kernels in the ultraviolet range by using a CCD Proline camera.

2.2.2 | UV/VIS–NIR Spectroscopy

A UV/VIS–NIR Spectrometer (Model T70 Plus, PG Instruments Ltd., United Kingdom) with a range of 100–1150 nm was applied to detect the fungi-infected wheat kernels. The spectroscopy detection technique is based on the principle that different chemical bonds in the tested sample absorb or emit different wavelengths of light when irradiated by a continuously changing frequency of UV/VIS–NIR light (Osborne, Fearn, and Hindle 1993). For spectroscopy methods, there are mainly three modes of data collection, namely, reflectance, transmittance and interaction (Wang et al. 2022). Determining a suitable measurement mode relies on the type of sample and the constituents to be tested. In the current study, the absorption spectra of samples were measured. Thin transparent sample tablets were prepared by mixing 10 mg of flour and 400 mg of ground potassium bromide (KBR) and pressing it in a die, forming a pellet at 10 tons and 1 min (Forato, Bernardes-Filho, and Colnago 1998; Sharma and Venugopalan 2014). To calibrate the device at the start of each stage, pure potassium bromide tablets were placed in a spectrometer. Since the absorption spectrum of potassium bromide in the wavelength range from 100 to 1150 nm was minimal, the entire spectrum of the sample was obtained.

2.3 | Data Processing and Model Development

2.3.1 | Image Processing

In numerous pattern recognition and computer vision tasks, leveraging color information proves invaluable for refining image analysis and optimizing segmentation outcomes, surpassing grayscale-based approaches. While the conventional color feature model relies on the primary spectral components of red (R), green (G), and blue (B), it falls short in scenarios where slight variations in light intensity significantly alter object descriptions. Hence, a color space less susceptible to such fluctuations is essential. Consequently, alternatives such as HSV, HSI, $L^*A^*B^*$, and others have gained prominence for comprehensive processing (Alipasandi, Ghaffari, and Alibeyglu 2013; Forato, Bernardes-Filho, and Colnago 1998). The methodology involved in this process can be outlined as follows:

a. The initial step involves acquiring images within the RGB color space.

b. Subsequently, the images are decomposed into their respective red (R), green (G), and blue (B) matrices. Following this, the mean of each matrix is computed across two dimensions.

c. Transitioning to the $L^*A^*B^*$ color space, individual calculations for L^*, A^*, and B^* values are conducted.

Because of the problem associated with mixing negative light, the Commission Internationale d'Eclairage (CIE) also developed a new color space called XYZ, which contains all the pure spectral colors within its positive octant. The transformation from RGB to XYZ is given by (The MathWorks Inc. 2011):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

The L^* component of lightness is defined as:

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) \quad (2)$$

where Y_n is the luminance value for nominal white (Fairchild 2013).

$$f(t) = \begin{cases} t^{1/3} & t > \delta^3 \\ t/\left(3\delta^2\right) + 2\delta/3 & else \end{cases} \quad (3)$$

$f(t)$ is a finite-slope approximation to the cube root with $\delta = 6.29$. The resulting 0–100 scale roughly measures equal amounts of lightness perceptibility.

In a similar fashion, the A^* and B^* components are defined as (The MathWorks Inc. 2011):

$$A^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \text{ and } B^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \quad (4)$$

where, X_n, Y_n, and Z_n is the measured white point (Szeliski 2010).

d. Transferring to HSV color space using the existing function in the MATLAB image processing Toolbox for this purpose and calculating means of H, S, and V matrices in two dimensions as follows:

First, the R, G, and B values are divided by 255 to change the range from 0.255 to 0.1 as follows (The MathWorks Inc. 2011):

$$R' = R/255$$
$$G' = G/255 \quad (5)$$
$$B' = B/255$$

$$C_{max} = max\left(R'.G'.B'\right)$$
$$C_{min} = min\left(R'.G'.B'\right) \quad (6)$$
$$\Delta = C_{max} - C_{min} G' = G/255$$

Then, the means of H, S, and V matrices were computed using the following equation (The MathWorks Inc. 2011):

$$H = \begin{cases} 0^{\circ}\,\Delta = 0 \\ 60^{\circ} \times \left(\frac{G'-B'}{\Delta} mod6\right). \quad C_{max} = R' \\ 60^{\circ} \times \left(\frac{B'-R'}{\Delta} + 2\right). \quad C_{max} = G' \\ 60^{\circ} \times \left(\frac{R'-G'}{\Delta} + 4\right). \quad C_{max} = B' \end{cases} \quad (7)$$

where mod6 refers to the modulo operation that gives the remainder when a number is divided by 6.

$$S = \begin{cases} 0. & C_{max} = 0 \\ \frac{\Delta}{C_{max}}. & C_{max} \neq 0 \end{cases} \quad (8)$$

$$V = C_{max} \quad (9)$$

e. Transferring to HSI color space using the function produced in the MATLAB M-File editor and calculating the mean of H, S, and I matrices in two dimensions.

Given an image in RGB color format, the H component of each RGB pixel is obtained using the following equation (The MathWorks Inc. 2011):

$$H = \begin{cases} \theta & if\, B \leq G \\ 360 - \theta & if\, B > G \end{cases} \quad (10)$$

which:

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\left[(R-G)^2 + (R-B)(G-B)\right]^{1/2}} \right\} \quad (11)$$

The saturation component is given by (The MathWorks Inc. 2011):

$$S = 1 - \frac{3}{(R+G+B)}[\min(R.G.B)] \quad (12)$$

In the end, the intensity component is determined by:

$$I = \frac{1}{3}(R+G+B) \quad (13)$$

It is assumed that the RGB values have been normalized to fall in the range of [0, 1] and the angle is measured concerning the red axis of the HSI space (Gonzalez 2007).

f. Likewise, employing the corresponding functions, YCbCr and YIQ color spaces were derived from the mean matrix of the R, G, and B values.

Finally, 18 color features were used as input for the neural network, as outlined in Table 1.

2.3.2 | Principal Component Analysis

PCA is one of the multivariate data analysis methods whose main purpose is to reduce the dimensions of studying problems (Shahin and Symons 2011; Wang, Dowell, and Lacey 1999; Zhang et al. 2007). The primary objective of applying PCA in this study was to reduce the dimensionality of the spectral data by identifying the most significant wavelengths that contribute to the variance in the dataset, thereby determining the most effective wavelengths for accurate sample classification. The Data Reduction Toolbox in IBM SPSS 16.0 software was used for PCA. KMO (Kaiser–Meyer–Olkin measure of sampling adequacy) and Bartlett's test were initially applied to investigate the ability of factor analysis on data (Kinnear and Gray 2006).

2.3.3 | Data Analysis Based on ANNs

The MLP algorithm stands out as a prevalent neural network (NN) topology extensively employed for classification tasks in agricultural products (Haykin 1998). However, in the realm of pattern recognition, feedforward networks take precedence, as they are adept at classifying inputs based on predefined target classes (Alipasandi et al. 2023). In this particular investigation, a feedforward network (FFN) featuring a single hidden layer with varying neuron configurations was implemented using the Neural Network Toolbox within MATLAB (R2011a). A pattern consisting of 18 components extracted from image processing was applied as an input in ANNs. According to Equation (14), different numbers of neurons were tested for the hidden layer and the proper number with Least MSE was chosen. Finally, the mean of classification was obtained in three replications (Demuth et al. 2014).

$$N_h = \frac{N_s}{(alpha \times (N_i + N_o))} \quad (14)$$

where, N_i is the number of input neurons, N_o is the number of output neurons, N_s is the number of samples in the training data set, N_h is the number of neurons for hidden layers, and alpha is an arbitrary scaling factor usually 2–10.

TABLE 1 | Features extracted from different color spaces.

Color spaces	RGB			HSI			*L*A*B**			XYZ			YCbCr			YIQ		
Features	R	G	B	H	S	I	L	A	B	X	Y	Z	Y	Cb	Cr	Y	I	Q

The network performance function was considered in "MSE" form, which measures the network's performance according to the mean of squared errors. The mean square of error (MSE) was calculated for various models to select the optimum network. The lower the error, the better the model performance is (Hornik 1991):

$$MSE = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}{n} \quad (15)$$

where $\hat{y}$ is the target and y is the observed output and n is the number of data sets (Blanco, Alcalà, and Sun 2009).

Typically, pattern-recognizing networks undergo supervised training, where each input corresponds to a specific target. In this study, the target consisted of a 6×300 matrix, representing six categories with 50 samples each for the UV/VIS–NIR imaging system. Similarly, in the spectroscopy method, inputs were derived from four components obtained through PCA. Various neuron configurations for the hidden layer were tested, with the optimal number determined based on the mean classification across three replications. The target in the UV/VIS–NIR spectroscopy method comprised a 3×42 matrix, representing three categories with 14 samples each. The Levenberg–Marquardt algorithm (trainlm) was utilized to train the ANN, with the Tangent sigmoid function applied in both the hidden and output layers.

During supervised training, the input and target data were randomly divided into training, validation, and test sets. Training proceeded until the network's performance plateaued on the validation data, indicating optimal generalization. The test data served as an independent assessment of the network's performance, ensuring robustness. The data split for training, validation, and test sets was 60%, 15%, and 25%, respectively.

2.3.4 | Correlation Feature Selection (CFS) Analysis

In this study, the CFS method was also used to select the wavelengths that provide the best classification among infected and healthy groups in which wavelengths are considered as features. This method evaluates the worth of a subset of features by considering each feature's individual predictive ability and the degree of redundancy between them. A feature is considered redundant if one or more of the other features are highly correlated with it. The CFS measure evaluates subsets of features based on the following hypothesis: "Good feature subsets contain features highly correlated with the classification, yet uncorrelated to each other. Subsets of features that are highly correlated with the class while having low intercorrelation are preferred" (Khan et al. 2021). A feature is considered good if it is more relevant to the class and not redundant to any other class features. For this, it uses information theory based on entropy. Entropy is a measure of the uncertainty of a random variable. It can be defined by the following equation (Quinlan 1993):

$$H(X) = -\sum P(x_i)\log_2\left(P(x_i)\right) \quad (16)$$

The entropy of X after observing values of another variable Y is defined in Equation (17) as follows (Quinlan 1993):

$$H(X|\,Y) = -\sum P(y_j)\sum P(x_i|\,y_j)\log_2\left(P(x_i|\,y_j)\right) \quad (17)$$

Here, $P(x_i)$ is the prior probabilities for all values of X, and $P(x_i|\,y_i)$ is the posterior probabilities of X when values of Y are given. The amount by which the entropy of X decreases reflects additional information about X provided by Y is called information gain given the equation $P(x_i|\,y_i)$ as:

$$IG(X|\,Y) = H(X) - H(X|\,Y) \quad (18)$$

It can be concluded that feature Y is regarded to be more correlated to feature X than to feature Z if $IG(X|\,Y) > IG(Z|\,Y)$. There is one more measure of symmetrical uncertainty which shows a correlation between features defined by Equation (19) as follows (Liu et al. 2002):

$$SU(X|\,Y) = 2\left[IG(X|\,Y)\,/\,(H(X) + H(Y))\right] \quad (19)$$

SU compensates information gain's bias toward features with more values and normalizes its value to a range of [0,1] with 1 showing that knowledge of either one completely predicts the value of the other and 0 showing that X and Y are independent. It considers a pair of features symmetrically. Entropy-based measures require nominal features, but they can be applied to measure correlations between continuous features as well if they are discretized properly.

To employ this method, weka version 3.8 was used to select features that enable a better categorization between classes and four wavelengths 230, 300, 380, and 1090 nm were found as dominant features. After selecting features that made good discrimination between classes, these features were applied to a C4.5 decision tree implemented in Weka. C4.5 (J48) is an algorithm implemented to generate a decision tree developed by Ross Quinlan (Salzberg 1994). C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, so C4.5 is often referred to as a statistical classifier.

3 | Results and Discussion

3.1 | Imaging Process

Considering the 18 color characteristics extracted from image processing as the neural network inputs and three replications of network training, the best result was obtained when a CCD Proline camera in the ultraviolet wavelength range was used. As seen in Table 2, it is evident that the total classification accuracy in the BPNN classifier for two kinds of wheat samples was 44.4% with a mean square error of 0.109297. Similarly, the percentages of classification by different cameras in visible and invisible ranges are reported in Table 3. As can be seen in Table 3, imaging systems have low accuracy using different digital color imaging systems for the Alvand and Sardari wheat classes. Overall, the BPNN classifier presented poor results in detecting healthy and infected wheat

TABLE 2 | Classification accuracy of Alvand and Sardari wheat cultivars using digital color imaging system (CCD camera in the range of ultraviolet wavelength).

Wheat sample	Training (%)	Validation (%)	Test (%)
Alvand			
Healthy	7.3	6.7	0.0
Infected with *Fusarium*	7.3	4.4	4.4
Infected with *Penicillium*	8.0	4.4	4.4
Sardari			
Healthy	8.0	6.7	15.6
Infected with *Fusarium*	7.3	6.7	8.9
Infected with *Penicillium*	9.3	4.4	11.1
Total	47.3	33.3	44.4

TABLE 3 | Total classification accuracies of Alvand and Sardari wheat classes using different digital color imaging systems.

No.	Camera	Optical ranges	Total classification accuracy (%)
1	CCD Proline	Visible	35.3
2	CCD Proline	Infrared	36.4
3	Analogue color camera "AGNI"	Ultraviolet	20.5
4	Analogue color camera "AGNI"	Visible	31.9
5	Analogue color camera "AGNI"	Infrared	38.5
6	LT Lutron Digital Instrument	Infrared	12.4

kernels. According to the results obtained for this method, the highest and lowest classification accuracies were acquired by using a CCD Proline camera with an optical range of infrared and LT Lutron Digital Instrument with values of 36.4% and 12.40%, respectively. Majumdar and Jayas (2000) used 23 morphological features to classify Canadian wheat kernels. In the current system by implementing the morphological feature, higher classification accuracy was obtained due to morphological differences of wheat varieties. However, the purpose of this study was to investigate color changes in fungal-infected wheat kernels. Therefore, the results of the morphological feature were not reported.

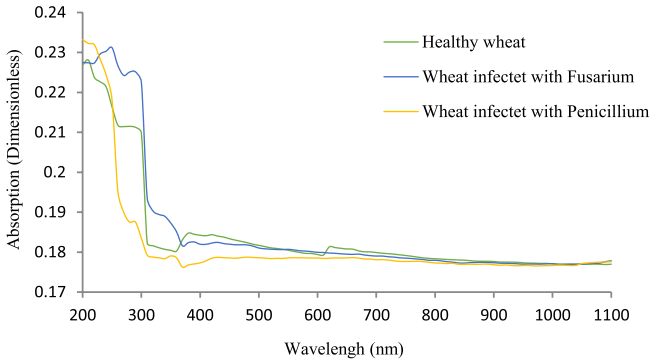

FIGURE 3 | Absorption spectra of healthy and fungus-infected wheat samples.

3.2 | Spectroscopy System

Figure 3 shows the absorption spectra of healthy and fungus-infected wheat samples acquired from Spectroscopy using GetData Graph Digitizer (version 2.3) so that wavelengths with a distance of 10 nm were used. The spectra of the wheat samples infected with *Fusarium* had a wider span of values when compared to those of the healthy wheat samples and wheat kernels infected with penicillium, which substantially overlap one another. In fact, the spectra of healthy samples were placed between two spectra of the fungus wheat kernels which makes the classification difficult.

KMO and Bartlett's tests were initially applied to investigate the ability of factor analysis on data. The response was acceptable for KMO and Bartlett's test in the wavelength range from 190 to 500 nm. The values for this range of wavelengths are outlined in Table 4.

KMO value of 0.676, as indicated in Table 4, indicates suitability for factor analysis. Additionally, Bartlett's test of sphericity yielded a significant result at the 1% level, implying a meaningful correlation between the variables and making data suitable for further exploration using factor analysis techniques.

Figure 4 shows the communalities and contributing factor of variables. Initial communality represents all communalities before extraction. Therefore, all of them are equal to one. As can be depicted in Figure 4, all communalities were more than 55.9%. These represent the ability of determined factors for variance explanation. However, there are some differences among the communalities. For instance, among communality for the wavelength of 360 and 400 nm which were 0.559 and 0.989, respectively.

Figure 5 shows the eigenvalues and the corresponding variances of factors associated with the first 10 components. The initial eigenvalues for each factor were estimated from the total explained variance. Explained variance is based on a percentage of the total variance and cumulative percentage. The eigenvalue for a given factor measures the variance in all variables which is accounted for by that factor. The ratio of eigenvalues is the ratio of explanatory importance of the factors concerning the variables. If a factor has a low eigenvalue, it signifies that it contributes minimally to the explanation of variances in the variables and may be ignored as redundant when compared to more important factors. The first four factors had eigenvalues greater

TABLE 4 | KMO and Bartlett's test.

Kaiser–Meyer–Olkin measure of sampling adequacy		0.676
Bartlett's test of sphericity	Approx. Chi-Square	3014
	Df[a]	351
	Sig[b]	0.000

[a]Degrees of freedom.
[b]Significance value.

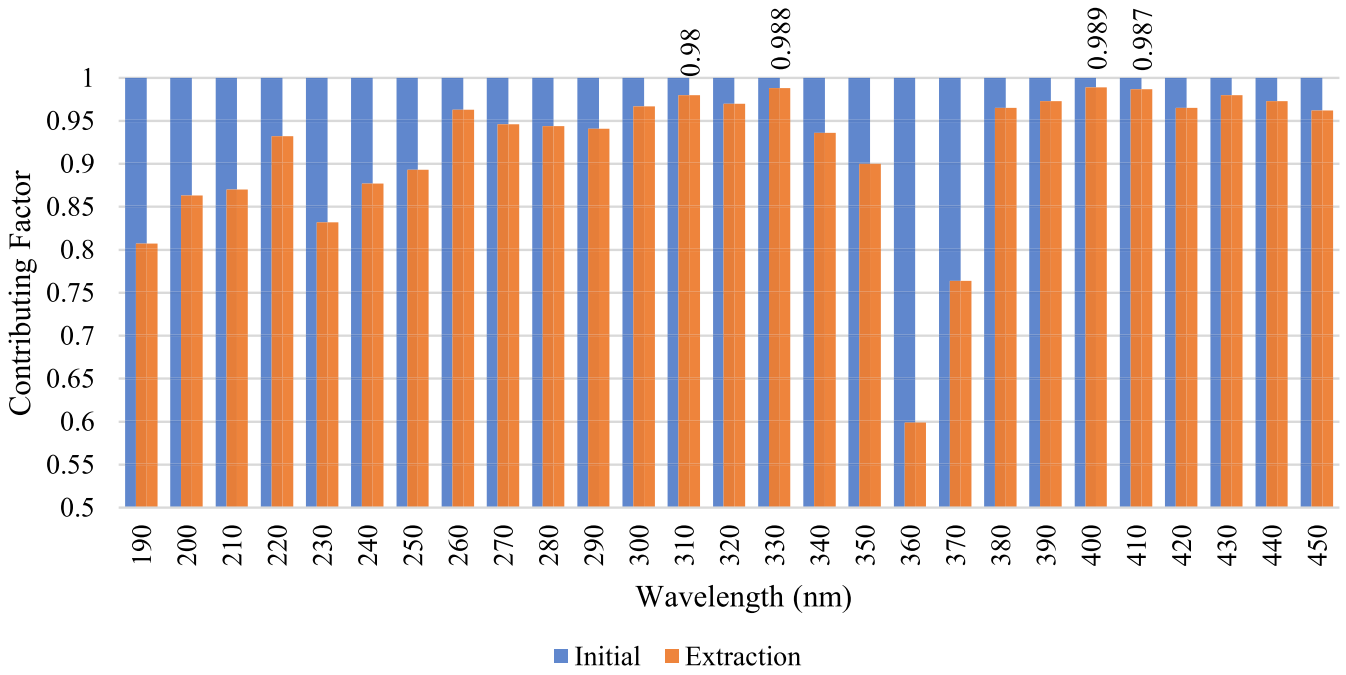

FIGURE 4 | Communalities and contributing factors of variables.

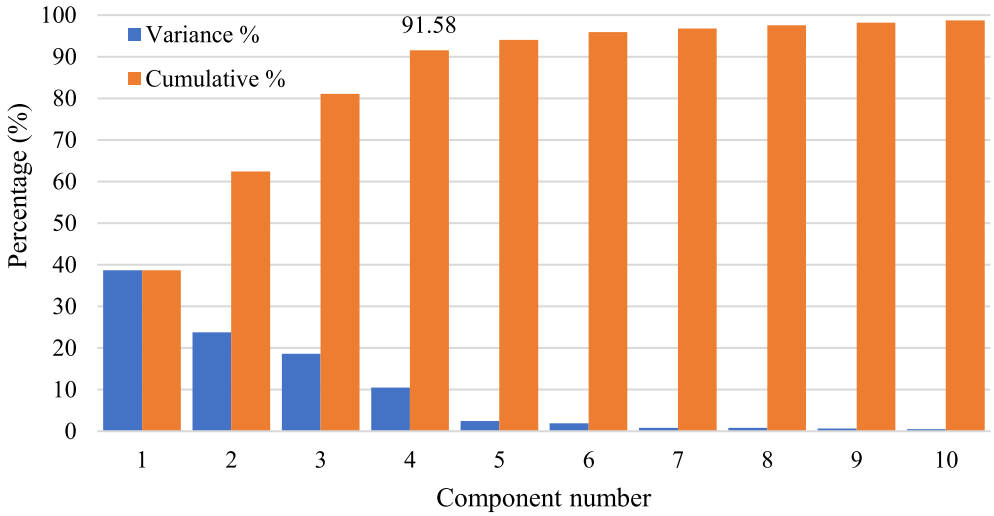

FIGURE 5 | Percentages of variance and cumulative of initial eigenvalues associated with the first 10 components.

than 1, as shown in Figure 5, cumulative variance for the first four factors is 91.58%.

Figure 6 shows the variation of eigenvalue against the factor number. This chart is used to determine the optimum number of components. It is observed that after the fourth factor, the change in eigenvalues is small. So, four factors can be extracted as important factors that have the greatest role in determining the variance.

Figure 7 demonstrates the contribution of variables. If squares of the factor loadings for each variable are summed, the extraction column of Figure 4 could be obtained. These coefficients explain the ability of determined factors to explain the variance of variables. Concurrently, they can be used to evaluate the suitability of variables for factor analysis.

As can be seen in Figure 7, wavelengths of 380, 390, 400, and 410 nm possess the greatest impact for the first factor. Similarly, wavelengths of 300, 310, 320, and 330 nm in the second factor, wavelengths of 230, 240, 250, and 260 nm in the third factor and wavelengths of 190, 200, 210, and 220 nm in the fourth factor presented the highest impact.

Thus, with regard to Figures 4 and 7, it can be concluded that the wavelengths 310, 330, 400, and 410 nm are the most influential wavelengths that have been used for classifying the samples by neural networks. Finally, by entering the following wavelengths to neural network input and training the network, 90.9% of wheat samples were correctly classified into healthy and unhealthy groups with a mean square error of 0.0185510 which is shown in Table 5.

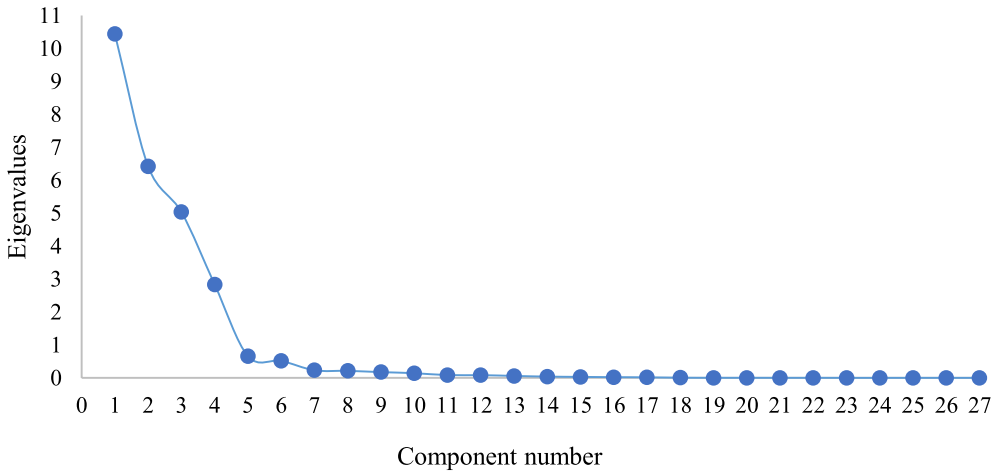

FIGURE 6 | A scree plot shows the eigenvalues on the *y*-axis and the number of factors on the *x*-axis.

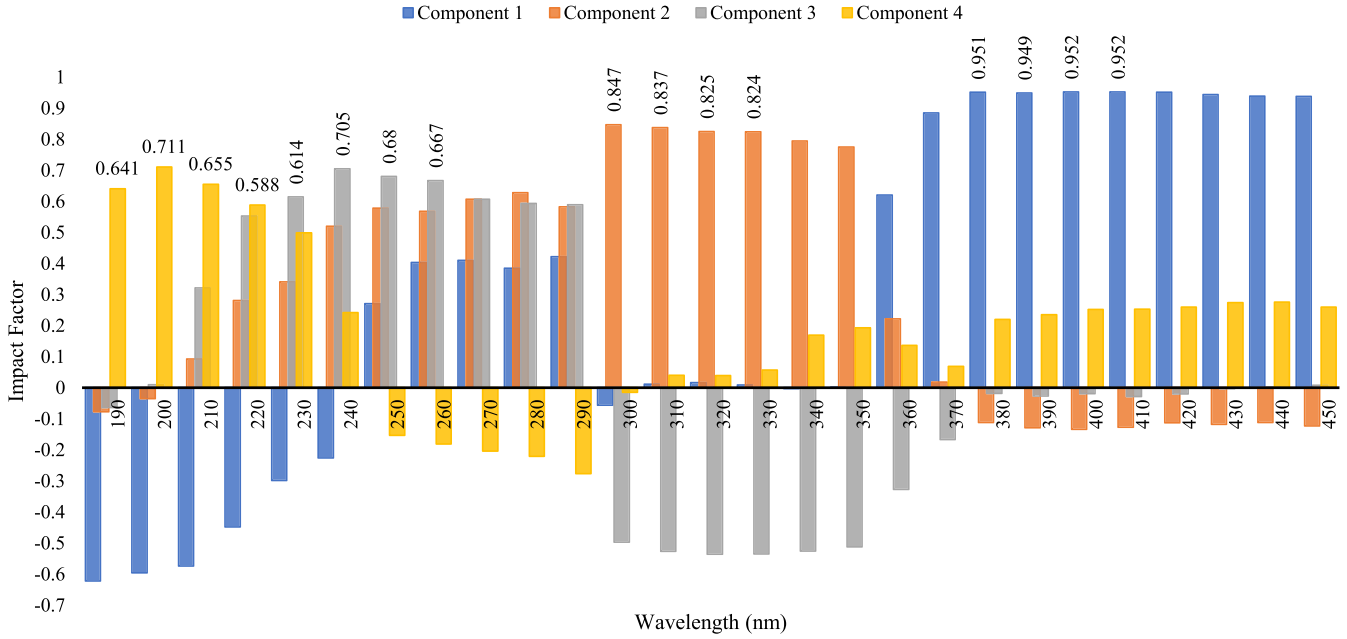

FIGURE 7 | Component matrix of variables, demonstrating the contribution of wavelengths.

TABLE 5 | Classification accuracies of wheat using spectroscopy methods for training, testing and validation.

Wheat sample	Training (%)	Validation (%)	Test (%)
Healthy wheat	100	100	100
Unhealthy wheat	89.5	100	88.9
Total	92.0	100	90.9

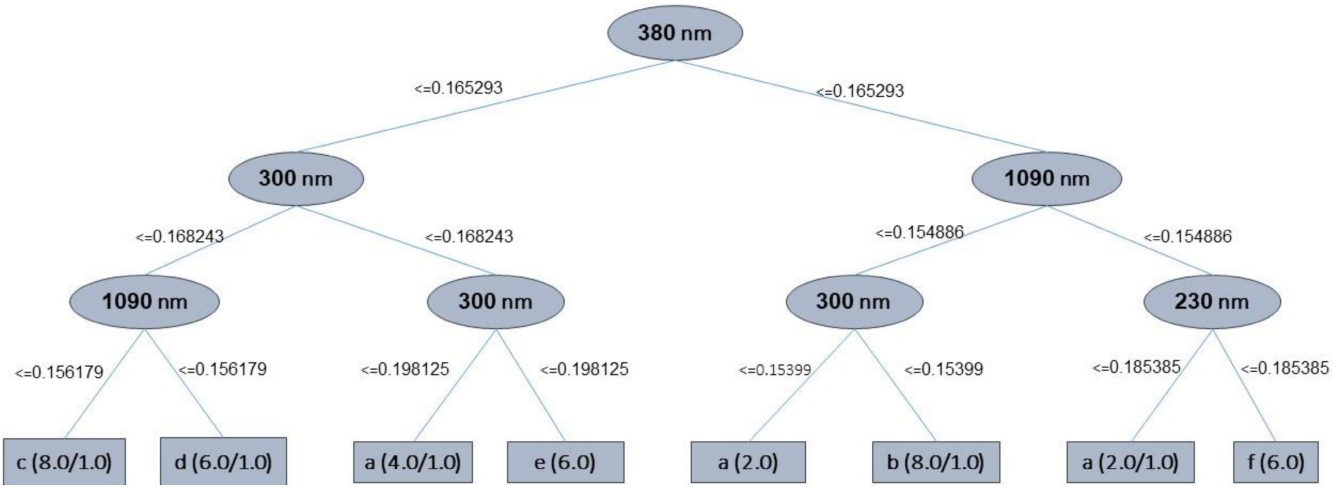

FIGURE 8 | Visualized decision tree used for categorizing sound and infected wheat kernels using selected features by the CFS method. a = Alvand, b = Alvand infected by *Penicillium*, c = Alvand infected by *Fusarium*, d = Sardari, e = Sardari infected by *Penicillium*, f = Sardari infected by *Fusarium*.

TABLE 6 | Confusion matrix of the decision tree classifier for selected features by the CFS method.

	a	b	c	d	e	f
a	6	0	0	1	0	0
	14.28%	0%	0%	2.38%	0%	0%
b	0	7	0	0	0	0
	0%	16.6%	0%	0%	0%	0%
c	0	0	7	0	0	0
	0%	0%	16.6%	0%	0%	0%
d	1	0	1	5	0	0
	2.38%	0%	2.38%	11.9%	0%	0%
e	1	0	0	0	6	0
	2.38%	0%	0%	0%	14.28%	0%
f	0	1	0	0	0	6
	0%	2.38%	0%	0%	0%	14.28%

Abbreviations: a = Alvand, b = Alvand infected by Penicillium, c = Alvand infected by *Fusarium*, d = Sardari, e = Sardari infected by *Penicillium*, f = Sardari infected by *Fusarium*.

These results can be compared with a HSI system used for wheat kernels reported by Singh et al. (2007). They were able to detect 95.5% of infected and healthy grains using a wavelength between 1000 and 1600 nm. Also, Wang et al. (2004) reported mean classification accuracies of 83%–95% using near-infrared spectroscopy to classify different types of fungal damage in soybean. However, using spectroscopy in the ultraviolet range is a more appropriate method for detecting fungal infected wheat kernels since compared with other methods, it has greater accuracy. Moreover, these spectral bands can easily be implemented on online sorting machines for the detection of fungal infected wheat kernels. Furthermore, Zhang et al. (2007) classified healthy and fungal infected wheat

kernels using hyperspectral images in 20 wavelengths between 1000 and 1600 nm. The overall classification accuracy was reported to be 94.8%. It could be concluded that cameras in visible and invisible ranges without monochromators and filters were unsuitable for detecting infected wheat kernels.

3.3 | CFS Measure

Figure 8 shows a visualized tree used for categorizing sound and infected wheat kernels using selected features by the CFS method. According to Figure 8, wavelength 380 nm is the wavelength that contains by far the highest level of information and for this reason, it has been selected as the first split criteria. Table 6 shows the confusion matrix of this classification. The sum of elements on the main diagonal of the confusion matrix divided by the total number of samples represents the correct classification of each classifier. As seen in Table 6 only five samples of 42 samples were misclassified and 88.1% of samples were classified correctly.

Figure 9 shows a visualized tree used for categorizing sound and infected wheat kernels using selected features by the PCA method. According to this tree, wavelength 400 nm is the wavelength that contains the highest level of information. Table 7 shows the confusion matrix of this classification. As can be seen in this table, by using this method, 88.1%

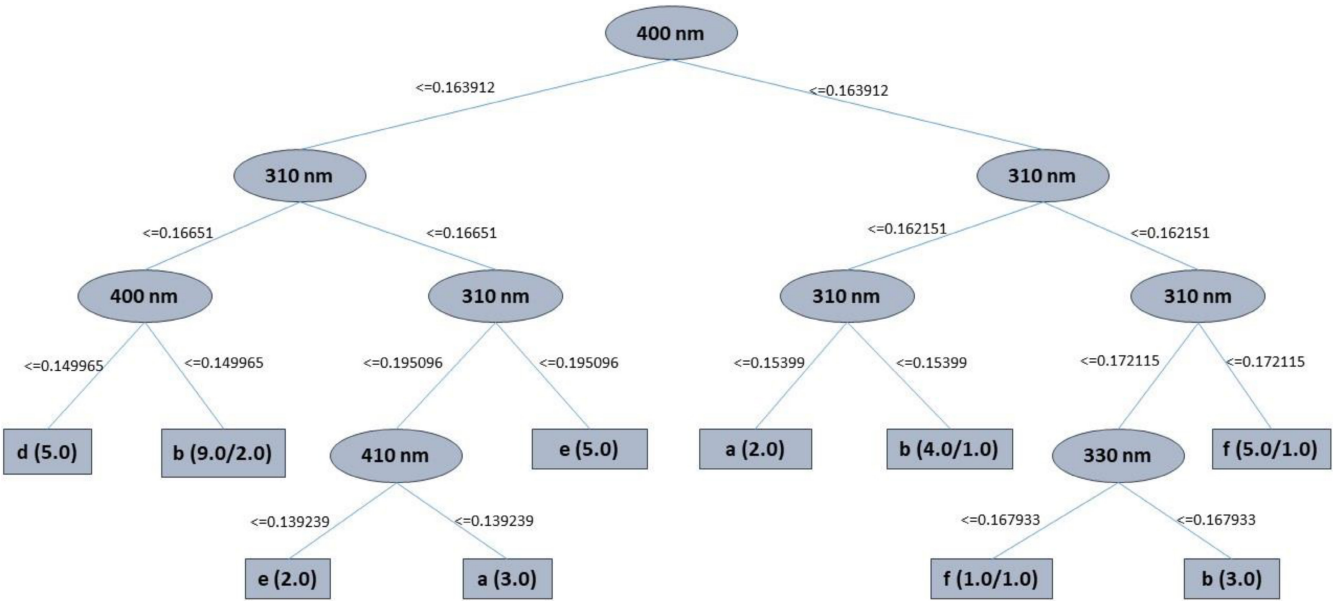

FIGURE 9 | Visualized decision tree used for categorizing sound and infected wheat kernels using selected features by PCA method. a = Alvand, b = Alvand infected by *Penicillium*, c = Alvand infected by *Fusarium*, d = Sardari, e = Sardari infected by *Penicillium*, f = Sardari infected by *Fusarium*.

TABLE 7 | Confusion matrix of the decision tree classifier selected features by PCA method.

	a	b	c	d	e	f
a	5	1	1	0	0	0
	11.9%	2.38%	2.38%	0%	0%	0%
b	0	6	0	0	0	1
	0%	14.28%	0%	0%	0%	2.38%
c	0	0	7	0	0	0
	0%	0%	16.6%	0%	0%	0%
d	0	0	1	5	0	1
	0%	0%	2.38%	11.9%	0%	2.38%
e	0	0	0	0	7	0
	0%	0%	0%	0%	16.6%	0%
f	0	0	0	0	0	7
	0%	0%	0%	0%	0%	16.6%

Abbreviations: a = Alvand, b = Alvand infected by Penicillium, c = Alvand infected by *Fusarium*, d = Sardari, e = Sardari infected by *Penicillium*, f = Sardari infected by *Fusarium*.

of samples were classified correctly, with only five out of 42 samples misclassified. Additionally, three classes were categorized perfectly.

4 | Conclusions

This study focuses on detecting and classifying fungal and healthy wheat kernels affected by *P. expansum* and *F. graminearum* using Spectroscopic technology and digital color imaging system through ANNs, PCA and FCM methodologies. In the spectroscopy system, four wavelengths (310, 330, 400, and 410 nm) corresponding to the PCA were found to be the most significant ones for optimum classification. The accuracy of this system was 90.9%, which makes it a suitable technique for separating healthy wheat kernels from kernels infected with *P. expansum* and *F. graminearum*. Among imaging systems, the best result was obtained using a CCD Proline camera in ultraviolet ranges with a detection accuracy of 44.4%. Results indicated that without special filters, it is very difficult to detect fungal-infected wheat kernels by imaging systems. The spectroscopy method gave better classification results than the imaging method. In this study, after the classification process for infected wheat kernels by CFS method, only five samples of 42 samples were misclassified and 88.1% of samples were correctly classified. Moreover, the results indicated that using cameras without monochromators or filters in visible and invisible ranges were unsuitable for detecting infected wheat kernels. The results obtained from this study show that the spectroscopy technology can potentially be used to automatically and rapidly detect infected wheat kernels with fungi. Moving forward into the future, the current study provides valuable insights into the intelligent determination of the infected wheat kernels with fungi through spectroscopy and developed methods. However, opportunities for improvement remain, including further enhancement of imaging systems through implanting advanced cameras and filters to ensure the universality of the method for different varieties of wheat and system enhancement for estimating the fungi concentrations in the spectroscopy methods.

Author Contributions

Saman Zohrabi: conceptualization, data curation, methodology, software, writing – original draft. **Seyed Sadegh Seiiedlou:** supervision, project administration, resources. **Iman Golpour:** software, visualization, writing – review – editing. **Mark Lefsrud:** validation, writing – review – editing. **Raquel P. F. Guiné:** writing – review – editing, formal analysis. **Barbara Sturm:** writing – review – editing, formal analysis, validation.

Acknowledgments

The authors would like to express their appreciation to the University of Tabriz for providing technical and financial support.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Alibeyglu, S. Z., H. Ghaffari, and A. Alipasandi. 2013. "Estimating Surface Area of Three Peach Varieties Using Allometric Relationships Obtain From Project Area." *International Journal of Agronomy and Plant Production* 4, no. 8: 1978–1984.

Alipasandi, A., A. Mahmoudi, B. Sturm, H. Behfar, and S. Zohrabi. 2023. "Application of Meta-Heuristic Feature Selection Method in Low-Cost Portable Device for Watermelon Classification Using Signal Processing Techniques." *Computers and Electronics in Agriculture* 205, no. August: 107578. https://doi.org/10.1016/j.compag.2022.107578.

Alipasandi, A., H. Ghaffari, and S. Z. Alibeyglu. 2013. "Classification of Three Varieties of Peach Fruit Using Artificial Neural Network Assisted With Image Processing Techniques." *International Journal of Agronomy and Plant Production* 4, no. 9: 2179–2186.

Almoujahed, M. B., A. K. Rangarajan, R. L. Whetton, et al. 2024. "Non-destructive Detection of Fusarium Head Blight in Wheat Kernels and Flour Using Visible Near-Infrared and Mid-Infrared Spectroscopy." *Chemometrics and Intelligent Laboratory Systems* 245: 105050. https://doi.org/10.1016/j.chemolab.2023.105050.

ASAE. 1988. "Moisture Measurement—Unground Grain and Seeds." *ASAE Standards* 1988, no. February: 567.

Beale, M. H., M. T. Hagan, and H. B. Demuth. 2012. "Neural Network ToolboxTM user's Guide. R2012a, The MathWorks, Inc., 3 Apple Hill Drive Natick, MA 01760–2098, Www. Mathworks. Com. Citeseer."

Blanco, M., M. Alcalà, and D. W. Sun. 2009. "Multivariate Calibration for Quantitative Analysis." In *Infrared Spectroscopy for Food Quality Analysis and Control*, 51–82. Dublin, IE: Elsevier.

Brdar, S., M. Panić, E. Hogeveen-van Echtelt, et al. 2021. "Predicting Sensitivity of Recently Harvested Tomatoes and Tomato Sepals to Future Fungal Infections." *Scientific Reports* 11, no. 1: 23109. https://doi.org/10.1038/s41598-021-02302-2.

Chavez, R. A., X. Cheng, T. J. Herrman, and M. J. Stasiewicz. 2022. "Single Kernel Aflatoxin and Fumonisin Contamination Distribution and Spectral Classification in Commercial Corn." *Food Control* 131: 108393. https://doi.org/10.1016/j.foodcont.2021.108393.

Chu, X., K. Zhang, H. Wei, et al. 2023. "A Vis/NIR Spectra-Based Approach for Identifying Bananas Infected With *Colletotrichum musae*." *Frontiers in Plant Science* 14: 1180203.

Dal-Pastro, F., P. Facco, F. Bezzo, E. Zamprogna, and M. Barolo. 2016. "Data-Driven Modeling of Milling and Sieving Operations in a Wheat Milling Process." *Food and Bioproducts Processing* 99: 99–108.

Demuth, H. B., M. H. Beale, O. de Jess, and M. T. Hagan. 2014. *Neural Network Design*. Martin Hagan.

Fairchild, M. D. 2013. *Color Appearance Models*. John Wiley & Sons.

FAO. 2021. "World Food and Agriculture—Statistical Yearbook 2021." In *World Food and Agriculture—Statistical Yearbook 2021*. Rome, Italy: FAO. https://doi.org/10.4060/cb4477en.

Femenias, A., M. B. Bainotti, F. Gatius, A. J. Ramos, and S. Marín. 2021. "Standardization of Near Infrared Hyperspectral Imaging for Wheat Single Kernel Sorting According to Deoxynivalenol Level." *Food Research International* 139: 109925. https://doi.org/10.1016/j.foodres.2020.109925.

Forato, L. A., R. Bernardes-Filho, and L. A. Colnago. 1998. "Protein Structure in KBr Pellets by Infrared Spectroscopy." *Analytical Biochemistry* 259, no. 1: 136–141. https://doi.org/10.1006/abio.1998.2599.

Gonzalez, T. F. 2007. "Handbook of Approximation Algorithms and Metaheuristics." In *Handbook of Approximation Algorithms and Metaheuristics*, 1–1432. New York: Chapman and Hall/CRC. https://doi.org/10.1201/9781420010749.

Haykin, S. 1998. *Neural Networks: A Comprehensive Foundation*. New Jersey: Prentice Hall PTR.

Hedayati, M. T., and R. A. Mohammadpour. 2005. "The Contamination Rate of Stored Wheat Samples of Mazandaran Province by *Aspergillus flavous* and Aflatoxin (2003)." *Journal of Kerman University of Medical Sciences* 9, no. 1: 52–61.

Hornik, K. 1991. "Approximation Capabilities of Multilayer Feedforward Networks." *Neural Networks* 4, no. 2: 251–257.

Jia, B., W. Wang, X. Z. Ni, X. Chu, S. C. Yoon, and K. C. Lawrence. 2020. "Detection of Mycotoxins and Toxigenic Fungi in Cereal Grains Using Vibrational Spectroscopic Techniques: A Review." *World Mycotoxin Journal* 13, no. 2: 163–177. 10.3920/WMJ2019.2510.

Khan, M. A., T. Akram, M. Sharif, M. Alhaisoni, T. Saba, and N. Nawaz. 2021. "A Probabilistic Segmentation and Entropy-Rank Correlation-Based Feature Selection Approach for the Recognition of Fruit Diseases." *EURASIP Journal on Image and Video Processing* 2021, no. 1: 1–28.

Kinnear, P. R., and C. D. Gray. 2006. "SPSS 12 Made Simple." In *SPSS 12 Made Simple*. London, UK: Psychology Press. https://doi.org/10.4324/9780203497388.

Liu, H., F. Hussain, C. L. Tan, and M. Dash. 2002. "Discretization: An Enabling Technique." *Data Mining and Knowledge Discovery* 6: 393–423.

Liu, Y., J. Zhang, H. Yuan, et al. 2022. "Non-Destructive Quality-Detection Techniques for Cereal Grains: A Systematic Review." *Agronomy* 12, no. 12: 3187.

Ma, H., H. Y. Ji, and W. S. Lee. 2016. "Identification of the Citrus Greening Disease Using Spectral and Textural Features Based on Hyperspectral Imaging." *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis* 36, no. 7: 2344–2350. https://doi.org/10.3964/j.issn.1000-0593(2016)07-2344-07.

Majumdar, S., and D. S. Jayas. 2000. "Classification of Cereal Grains Using Machine Vision: I. Morphology Models." *Transactions of the American Society of Agricultural Engineers* 43, no. 6: 1669–1675. https://doi.org/10.13031/2013.3107.

Mustafa, G., H. Zheng, I. H. Khan, et al. 2024. "Enhancing Fusarium Head Blight Detection in Wheat Crops Using Hyperspectral Indices and Machine Learning Classifiers." *Computers and Electronics in Agriculture* 218: 108663. https://doi.org/10.1016/j.compag.2024.108663.

Osborne, B. G., T. Fearn, and P. H. Hindle. 1993. *Practical NIR Spectroscopy With Applications in Food and Beverage Analysis*. 2nd ed., 227. Essex, UK: Harlow and New York: Longman Scientific & Technical, Wiley.

Parrag, V., Z. Gillay, Z. Kovács, et al. 2020. "Application of Hyperspectral Imaging to Detect Toxigenic Fusarium Infection on Cornmeal." *Progress in Agricultural Engineering Sciences* 16, no. 1: 51–60.

Pearson, T. C., and D. T. Wicklow. 2006. "Detection of Corn Kernels Infected by Fungi." *Transactions of the ASABE* 49, no. 4: 1235–1245. https://doi.org/10.13031/2013.21723.

Peiris, K. H. S., M. O. Pumphrey, and F. E. Dowell. 2009. "NIR Absorbance Characteristics of Deoxynivalenol and of Sound and Fusarium-Damaged Wheat Kernels." *Journal of Near Infrared Spectroscopy* 17, no. 4: 213–221. https://doi.org/10.1255/jnirs.846.

Qin, J. 2010. "Hyperspectral Imaging Instruments." In *Hyperspectral Imaging for Food Quality Analysis and Control*, edited by D.-W. Sun, 129–172. San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-374753-2.10005-X.

Quinlan, J. R. 1993. "Program for Machine Learning. C4.5."

Salzberg, S. L. 1994. "C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993." *Machine Learning* 16, no. 3: 235–240. https://doi.org/10.1007/BF00993309.

Schmidt, M., S. Horstmann, L. de Colli, et al. 2016. "Impact of Fungal Contamination of Wheat on Grain Quality Criteria." *Journal of Cereal Science* 69: 95–103.

Sha, W., K. Hu, and S. Weng. 2023. "Statistic and Network Features of RGB and Hyperspectral Imaging for Determination of Black Root Mold Infection in Apples." *Food* 12, no. 8: 1608. https://doi.org/10.3390/foods12081608.

Shahin, M. A., and S. J. Symons. 2011. "Detection of Fusarium Damaged Kernels in Canada Western Red Spring Wheat Using Visible/Near-Infrared Hyperspectral Imaging and Principal Component Analysis." *Computers and Electronics in Agriculture* 75, no. 1: 107–112. 10.1016/j.compag.2010.10.004.

Sharma, B., and K. Venugopalan. 2014. "Comparison of Neural Network Training Functions for Hematoma Classification in Brain CT Images." *IOSR Journal of Computer Engineering* 16, no. 1: 31–35. https://doi.org/10.9790/0661-16123135.

Singh, C. B., D. S. Jayas, J. Paliwal, and N. D. G. White. 2007. "Fungal Detection in Wheat Using Near-Infrared Hyperspectral Imaging." *Transactions of the ASABE* 50, no. 6: 2171–2176.

Singh, C. B., D. S. Jayas, J. Paliwal, and N. D. G. White. 2012. "Fungal Damage Detection in Wheat Using Short-Wave Near-Infrared Hyperspectral and Digital Colour Imaging." *International Journal of Food Properties* 15, no. 1: 11–24.

Szeliski, R. 2010. *Computer Vision: Algorithms and Applications*. Berlin: Springer Science & Business Media.

The MathWorks Inc. 2011. "Image Processing Toolbox."

Visen, N. S., J. Paliwal, D. S. Jayas, and N. D. G. White. 2004. "Image Analysis of Bulk Grain Samples Using Neural Networks." *Canadian Biosystems Engineering/Le Genie Des Biosystems Au Canada* 46: 7.11–7.15. https://doi.org/10.13031/2013.15002.

Wagacha, J. M., and J. W. Muthomi. 2008. "Mycotoxin Problem in Africa: Current Status, Implications to Food Safety and Health and Possible Management Strategies." *International Journal of Food Microbiology* 124, no. 1: 1–12. https://doi.org/10.1016/j.ijfoodmicro.2008.01.008.

Wang, D., F. E. Dowell, and R. E. Lacey. 1999. "Single Wheat Kernel Color Classification Using Neural Networks." *Transactions of the American Society of Agricultural Engineers* 42, no. 1: 233–240. https://doi.org/10.13031/2013.13200.

Wang, D., F. E. Dowell, M. S. Ram, and W. T. Schapaugh. 2004. "Classification of Fungal-Damaged Soybean Seeds Using Near-Infrared Spectroscopy." *International Journal of Food Properties* 7, no. 1: 75–82. https://doi.org/10.1081/JFP-120022981.

Wang, M., Y. Xu, Y. Yang, B. Mu, M. A. Nikitina, and X. Xiao. 2022. "Vis/NIR Optical Biosensors Applications for Fruit Monitoring." *Biosensors and Bioelectronics: X* 11: 100197. https://doi.org/10.1016/j.biosx.2022.100197.

Wang, N., N. Zhang, F. E. Dowell, and T. Pearson. 2005. "Determining Vitreousness of Durum Wheat Using Transmitted and Reflected Images." *Transactions of the American Society of Agricultural Engineers* 48, no. 1: 219–222. https://doi.org/10.13031/2013.17920.

Xin, Z., L. Wang, C. Liu, et al. 2023. "Detection of Infection Degree of Botrytis Disease in Apple by Fusing Spectra and Image Features From Hyperspectral Imaging." https://doi.org/10.21203/rs.3.rs-2623242/v1.

Yang, G., Q. Wang, C. Liu, X. Wang, S. Fan, and W. Huang. 2018. "Rapid and Visual Detection of the Main Chemical Compositions in Maize Seeds Based on Raman Hyperspectral Imaging." *Spectrochimica Acta—Part A: Molecular and Biomolecular Spectroscopy* 200: 186–194. https://doi.org/10.1016/j.saa.2018.04.026.

Zhang, H., J. Paliwal, D. S. Jayas, and N. D. G. White. 2007. "Classification of Fungal Infected Wheat Kernels Using Near-Infrared Reflectance Hyperspectral Imaging and Support Vector Machine." *Transactions of the ASABE* 50, no. 5: 1779–1785. https://doi.org/10.13031/2013.23935.

Zhao, T., M. Chen, X. Jiang, et al. 2020. "Integration of Spectra and Image Features of Vis/NIR Hyperspectral Imaging for Prediction of Deoxynivalenol Contamination in Whole Wheat Flour." *Infrared Physics and Technology* 109: 103426. https://doi.org/10.1016/j.infrared.2020.103426.

Zohrabi, S., S. S. Seiiedlou, E. Safari, and Y. Hampanejad. 2015. "Spectral Reaction Investigation of Healthy and Fungi Infected Wheat Kernels Using Imaging in Visible and Non-visible Range and Utilization of Neural Network." *Journal of Researches in Mechanics of Agricultural Machinery* 3, no. 2: 1–9. http://journals.sku.ac.ir/page/article-frame.html?articleId=237747&langId=fa.

Zohrabi, S., S. S. Seiiedlou, Y. Hampanejad, and A. Aalipasandi. 2015. "Detection of Fungal Infected Wheat Kernels Using Spectroscopy Method and Utilization of ANFIS System." In *9th National Congress on Agricultural Machinery Engineering (Biosystems and Mechanization)*, 102–110. Tehran University, Karaj, Iran: University College of Agriculture and Natural Resources.