Research Paper

Detecting Fusarium head blight in wheat kernels using hyperspectral imaging

Jayme G.A. Barbedo a,*, Casiane S. Tibola b,1, José M.C. Fernandes b,1

a Embrapa Agricultural Informatics, Av. André Tosello, 209, C.P. 6041, Campinas 13083-886, SP, Brazil
b Embrapa Wheat, Rodovia BR-285, Km 294, C.P. 3081, Passo Fundo 99001-970, RS, Brazil

Article info

Article history:
Received 3 November 2014
Received in revised form 9 January 2015
Accepted 20 January 2015
Published online 5 February 2015

Keywords:
Fusarium index
Fusarium-damaged wheat kernels
Image processing

Because of the health risks associated with the ingestion of the mycotoxin deoxynivalenol (DON) produced by Fusarium head blight (FHB), improving its detection in wheat kernels is a major research goal. Currently, assessments are largely performed visually by human experts. Being subjective, such assessments may not always be consistent or entirely reliable. As a result, methods with a higher degree of objectivity have been investigated, and special attention has been dedicated to the use of hyperspectral imaging (HSI) as the basis for more reliable detection strategies. This paper presents an algorithm for automatic detection of FHB in wheat kernels using HSI. The goal was to develop a simple and accurate algorithm which gave as output an index that can be interpreted as the likelihood of the kernel being infected by FHB. With a classification accuracy above 91%, the developed algorithm was robust to factors such as shape, orientation, shadowing and clustering of kernels. It was shown that the algorithm was not only suitable for detecting FHB, but it also has the capability, albeit limited, of estimating DON concentrations in wheat kernels.

© 2015 IAgrE. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Wheat is one of the main crops grown in the world, being important both socially and economically. As such, considerable effort is placed on the control of diseases that affect production. The Fusarium head blight (FHB), also known as scab and caused by the fungal plant pathogen Fusarium graminearum (Gibberella zeae), is particularly devastating (Schmale & Bergstrom, 2003). The disease poses a significant threat to both humans and animals since G. zeae produces the mycotoxin deoxynivalenol (DON) which can disrupt normal cell function by inhibiting protein synthesis (Schmale & Bergstrom, 2003). Thus, in order to avoid potential health risks, diseased grains must be identified before they are processed to avoid them being incorporated into food for humans and animal feed.

Normally, the detection of Fusarium head blight (FHB) is carried out manually by human experts using a process that may be both lengthy and tiresome. Moreover, the effectiveness of this kind of detection may drop with factors such as fatigue, external distractions and optical illusions. A study on the effectiveness of visual disease detection and classification...
was carried out by Bock, Poole, Parker, and Gottwald (2010). Thus, methods capable of performing this detection automatically are in high demand. Most automatic methods proposed to date rely on image processing to perform their tasks. The image-based detection of FHB can follow two strategies, one using images of wheat ears captured in the field, and the other using images of kernels usually captured in the laboratory under controlled conditions. Near-infrared (NIR), chlorophyll fluorescence (CF) and visible range (VR) methods are usually employed to obtain images. According to Bauriegel and Herppich (2014), the visible range, particularly in the orange and red bands (600–750 nm), is useful to detect FHB in wheat only between mid-grain development and early grain ripening, when the chlorophyll and water contents are still high. In later stages of grain development, only NIR, especially in the 1000–1400 nm wavelength range, has the potential to detect the stresses caused by the disease. Indeed, most methods proposed in the literature operate within those bands. An in-depth discussion and some examples about the most common imaging techniques and wavelengths can be found in Sankaran, Mishra, Ehsani, and Davis (2010) and in Bauriegel and Herppich (2014). A brief discussion on some of the main FHB detection techniques proposed is presented next.

One of the first studies on the use of hyperspectral imaging for detecting FHB in wheat kernels was carried out by Delwiche and Kim (2000). They used a custom-made system to gather hyperspectral images from kernels of three different wheat varieties, and used a ratio between bands to perform the separation between sound and diseased kernels. They concluded that hyperspectral imaging may be used for FHB detection, and that factors such as wheat variety and position of the kernels may affect the results.

Polder, van der Heijden, Waalwijk, and Young (2005) compared the effectiveness of using images captured in the visible and near-infrared ranges for detecting FHB in wheat kernels. In their analysis, they used techniques like relation between two wavelength bands, unsupervised fuzzy c-means clustering and supervised partial least squares regression. They concluded that the NIR range is more appropriate for the task.

Wegulo and Dowell (2008) conducted a study in which NIR and visual sorting of Fusarium-damaged kernels (FDK) were compared. They came to the conclusion that the tested system, besides correlating well with the visual measurements, had the advantage of being faster and more consistent than human operators.

Menesatti et al. (2009) used visible range images to perform a morphometric analysis of individual kernels. They employed Sobel filtering for edge detection and elliptic Fourier analysis for extracting shape information, which were then submitted to a partial least squares discriminant analysis. The accuracy of the proposed system was close to 70%.

Delwiche, Kim, and Dong (2011) presented a study on the assessment of Fusarium damage in wheat kernels, using NIR and VR images. Using linear discriminant analysis (LDA), they were able to achieve an accuracy of 95% with respect to the visual assessment, both with NIR and VR images. They also came to the conclusion that considering the entire kernel images yield better results than taking either a portion of the endosperm or the germ tip separately.

Shahin and Symons (2011) used NIR and VR images in the 400–1000 nm range, combined with principal component analysis (PCA) and LDA, to classify individual Canada Western Red Spring wheat kernels into sound, mildly damaged, and severely damaged by Fusarium. They concluded that NIR images can be used to detect FHB in that variety of wheat, and that six wavelengths can produce classifications as good as the entire NIR range.

As mentioned before, some studies have focused on the early detection of FHB by analysing images of wheat ears in the field, i.e. before harvesting. Bauriegel, Giebel, and Herppich (2010) investigated the potential of chlorophyll fluorescence imaging for detecting FHB in wheat ears, by means of the so-called potential maximum photochemical efficiency. Dammer, Möller, Rodemann, and Heppner (2011) used conventional and multispectral (red, infrared and green) cameras for detecting FHB in winter wheat ears; they applied conventional binarisation and area thresholding to each of the bands in order to isolate regions potentially infected. Finally, Bauriegel, Giebel, Geyer, Schmidt, and Herppich (2011) used hyperspectral imaging, combined with PCA, for detecting FHB in wheat ears before harvest; the detection was more successful in the beginning of the medium milk stage, and failed completely just after the start of flowering and in the fully ripe stage.

Despite the advances achieved so far, there is still much to be done. In this context, this paper has three main objectives. The first is to propose a new algorithm for automatic FHB detection using NIR images that is accurate, simple to implement, and computationally inexpensive, making it suitable for being used whenever NIR images are available. The second is to offer more insight into some of the factors that have an impact on the use of NIR images for FHB detection, such as kernel shape and orientation, specular reflections, spectral band combination, kernel shadowing and clustering and spatial resolution. The final objective is to investigate the viability of using the algorithm for estimating deoxynivalenol (DON) concentrations in wheat kernels.

The algorithm proposed here focuses on the detection of FHB in individual wheat kernels, using hyperspectral images captured in the 528–1785 nm wavelength range. The output of the algorithm, the so-called Fusarium index (FI), is a number

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>chlorophyll fluorescence</td>
</tr>
<tr>
<td>DON</td>
<td>deoxynivalenol</td>
</tr>
<tr>
<td>FHB</td>
<td>Fusarium head blight</td>
</tr>
<tr>
<td>FDK</td>
<td>Fusarium damaged kernels</td>
</tr>
<tr>
<td>FI</td>
<td>Fusarium index</td>
</tr>
<tr>
<td>HSI</td>
<td>hyperspectral imaging</td>
</tr>
<tr>
<td>LDA</td>
<td>linear discriminant analysis</td>
</tr>
<tr>
<td>NIR</td>
<td>near infrared</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PDF</td>
<td>probability density function</td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest</td>
</tr>
<tr>
<td>VR</td>
<td>visible range</td>
</tr>
</tbody>
</table>

DON: deoxynivalenol
CF: chlorophyll fluorescence
FHB: Fusarium head blight
FDK: Fusarium damaged kernels
FI: Fusarium index
HSI: hyperspectral imaging
LDA: linear discriminant analysis
NIR: near infrared
PCA: principal component analysis
PDF: probability density function
ROI: region of interest
VR: visible range
between zero and one that reveals the likelihood of a kernel to be infected with FHB. This value can also be used, to a certain extent, to estimate DON concentration levels, as it will be seen in Section 3.

Although the principles explored here are basically the same of previous studies using NIR hyperspectral imaging, the proposed algorithm has some advantages that are not always found in its predecessors:

- It is very simple to implement, employing only widely used morphological operations to achieve the desired results, and using only one band for the disease detection.
- It is computationally fast, as it does not rely on any complex technique to work properly.
- It does not require any manual tuning.
- It is capable of successfully separating clustered kernels without employing morphological operations such as erosion, dilation and opening, thus preserving kernel shapes.
- An implementation of the algorithm with a graphical interface is available on the internet for download at https://www.digipathos.cnptia.embrapa.br.

2. Material and methods

2.1. Image acquisition

The acquisition of the hyperspectral images was carried out at the Wheat Quality Laboratory of Embrapa Wheat, located in Passo Fundo, Brazil. The hyperspectral imaging (HSI) system used in this work is shown in Fig. 1, and is largely based on the guidelines presented in Delwiche, Souza, and Kim (2013), where more details are available. The spectrometer used in the system (EV/NIR Hyperspec Model 1003B-10151, Headwall Photonics Inc., MA, USA) employs an InGaAs sensor with a 320 x 256 pixels focal plane array (FPA) and a XENICS camera (Model XEVA-1246 XC 134, Leuven, Belgium); it measures the reflectance of the samples in the 528–1785 nm wavelength range (VR/NIR), with a spectral resolution of 5–7 nm, resulting in 256 bands. This spectrometer is coupled with a 25 mm C-mount lens (F1.35/25 mm) placed 440 mm above the samples, resulting in a line width field of view of 50 mm. The images are acquired in a line-by-line basis (push-broom acquisition) – a total of 800 lines are scanned for each sample, with approximately 500 of those lines delimiting the region of interest where the kernels are located. The illumination is provided by an external Quartz Tungsten-Halogen lamp, whose light is conveyed via an optical fibre bundle that terminates in a 250-mm long line tilted by approximately 30° with respect to the vertical axis, as it can be seen in Fig. 1e. With this setup, no diffuser is necessary, as the light is scattered by the use of the fibre bundle.

The images used in this work were captured using the setup described above, according to the following steps. Firstly, the wheat kernels were placed in a Teflon holder and transported to the conveyor unit (Fig. 1d). The conveyor speed, field of view (FOV), motor speed, exposure time, and wavelength were controlled by means of the Hyperspec III software (Headwall Photonics Inc., MA, USA). The images, acquired as the samples moved through the FOV, were post-processed in order to compensate for external interferences and to keep only the NIR response of the samples. This compensation was carried out by first acquiring white and dark reference images. The white reference (W) was obtained using a white Teflon material with a 99% reflectance, and the dark reference (D) was acquired by covering the lens with a zero-reflectance lens cover. These reference images were then used to calculate the actual reflectance value, according to:

\[
R = \frac{(R_0 - D_i)}{(W_i - D_i)}
\]

where \(R_0\) is the raw hyperspectral image, \(W\) is the white reference, \(D\) is the dark reference, \(i\) is the pixel index (\(i = 1,2,3, \ldots, n\)), and \(n\) is the total number of pixels.

The final result for each given sample is a 3D matrix with dimensions of 320 x 800 x 256, that is, the images have a spatial dimension of 320 x 800 over 256 different bands. It is important to mention that the data representation is “16-bit unsigned integer” prior to the compensation, and “32-bit single precision floating point” afterward. This is important because it affects some of the operations performed throughout the algorithm to be described next.

2.2. Image dataset

The database used in this work was composed by 27 hyperspectral images (803 kernels in total). Each image contained a mixture of 25–50 kernels of four varieties: BRS 194 (FHB susceptible, kernel with rounder shape); Quartzo (moderately susceptible to FHB, grain with rounder shape); BRS Parrudo and BRS 179 (both moderately resistant to FHB and with elongated shape). One third of the images had only sound kernels, one third had only diseased kernels, and the remaining images contain both diseased and sound kernels in approximately equal proportions.
The wheat samples were obtained from a greenhouse experiment to characterise the aggressiveness and mycotoxin production of *F. graminearum sensu stricto* (Fgss) and *Fusarium meridionale* (Fmer). The four wheat cultivars cited above were inoculated either by point-inoculation in the central floret or by spraying during the anthesis stage.

The wheat spikes were harvested when the grains had reached the maturity point. The dried ears were threshed in a stationary thresher (Precision Machine Company Inc., Lincoln, NE, USA). Sound and diseased kernels were then separated by visual inspection, in which an experienced technician would look for signs such as wrinkling, deformations, and whitish or pinkish colour on the kernel surface for identification of the *Fusarium* damaged kernels (FDK). The severity of the symptoms was not graded, and the separation of the kernels was performed by the same technician throughout the experiments.

DON concentrations were determined for the kernels present in six of the hyperspectral images in the database, in order to investigate the possibility of using the proposed algorithm for estimating DON levels in wheat kernels. DON concentrations were measured by liquid chromatography–mass spectrometry (LC–MS). The range of measurement was 250–5000 ppb, and the detection limit was 200 ppb. Each image contained, in average, 50 kernels with approximately the same DON concentration levels, as shown in Table 1.

### Algorithm

Figure 2 shows a workflow diagram that summarises the structure of the algorithm. Each block in the figure will be explained in detail in this section. As it can be seen, the only input the algorithm receives is the hyperspectral image. Figure 3 shows, as an example, a greyscale representation of the reflectance of a specific band (647 nm) of one image in the database.

#### Delimitation of the region of interest

The first step of the algorithm is to determine the region of interest (ROI) where the kernels are located. It was determined empirically that the reflectance in band 25 (647 nm, the same shown in Fig. 3) provides the best contrast between ROI and the background. Instead, four bands were selected: 672 nm (A), 1361 nm (B), 1509 nm (C), and 1657 nm (D), as shown in Fig. 6a–d. The greyscale representation of the reflectances in those bands were normalised as described earlier in this section. Then, the following operations were applied:

\[
\begin{align*}
dif_1 & = \frac{|A - D|}{\text{max}(|A - D|)} \quad (2) \\
dif_2 & = \frac{|B - C|}{\text{max}(|B - C|)} \quad (3) \\
\text{ref} & = 0.5 \cdot (dif_1 + dif_2) \quad (4)
\end{align*}
\]

where \(\text{dif}_1\), is the normalised difference between bands A and D (Fig. 6e), \(\text{dif}_2\), is the normalised difference between bands B and C (Fig. 6f), and \(\text{ref}\) is the combination of both normalised differences (Fig. 6g). It can be seen in Fig. 6 that, although \(\text{dif}_1\) provides a good separation between kernels and background, there are some parts of the kernels that actually appear very dark, hence the need for a combination with \(\text{dif}_2\). A threshold was then applied to \(\text{ref}\), where pixels were made equal to one if their values were larger than 0.15, and equal to zero otherwise. The mask for the kernels is shown in Fig. 6h. As it can be seen, a few spurious elements may remain. In the case of the example shown in Fig. 6, two elements must be discarded: a small object resulting from noise in the image, and a reference ruler located in the upper-middle section of the ROI.

For that, two simple rules were applied:

- If a given object was smaller than 0.0004 \(xp \times yp\) pixels, where \(xp\) and \(yp\) are the spatial dimensions of the image in pixels, it was discarded.
- If an object was larger than 0.006 \(xp \times yp\) pixels, and its area divided by the area of its convex hull was larger than 0.9, it was considered a foreign element (in this case, a ruler), and it was discarded.

#### Cluster splitting

Depending on the position the kernels, some may appear merged in the mask image (Fig. 7a), which means that they will...
be treated as a single entity by the algorithm. Thus, in order to avoid incorrect results, these objects must be split. The strategy normally used is to apply morphological operations such as erosion, dilation or opening. The problem with this approach is that those operations may distort the shape of the objects. Moreover, depending on the extension of the merge, those operations may remove entire kernels before separating them. The strategy used here preserves kernel shapes and is robust to most conditions, as described in the following.

When kernels touch, the nearly perfect convexity of isolated kernels is usually lost, as shown in Fig. 7a. If the convex hull (smallest convex region capable of encompassing a given object) is calculated for the mask corresponding to single kernel, the result will be either the original mask itself or something very close to it. On the other hand, if the convex hull is calculated for a cluster of kernels, the resulting object will be usually quite different from the original mask generated by that cluster. Such a difference can be clearly seen by comparing Fig. 7a and b. Also, when the original mask is subtracted from the convex hull, the result will be an image containing a number of objects (here called concave regions) that is directly proportional to the number of kernels in the cluster (Fig. 7c). Only concave regions containing ten pixels or more are considered, because smaller regions are usually the result of irregularities and/or noise in the original image. An interesting property of those regions is that the shortest line connecting a pair of them is usually the one that best separates the two kernels associated to that pair. This can be clearly visualised in Fig. 7c, where the gap between the large concave region and the lower region defines the best separation between the two bottom kernels. The application of this idea to clusters of two kernels is straightforward, as they will generate only two concave regions (one at either side of the touching point), making the separation only a matter of finding the shortest line that connect those two regions. However, higher order clusters usually generate more than two concave regions (as is the case in Fig. 7), making the process more complicated:

- When three concave regions are present, as is the case of Fig. 7, this indicates that there are three kernels in that cluster. In order to determine the lines that best separate those kernels, the shortest lines connecting all three possible pairs of regions are determined. However, only

![Fig. 2 - Workflow diagram of the basic structure of the algorithm.](image)

![Fig. 3 - Greyscale representation of the 647-nm band.](image)

![Fig. 4 - ROI determination, a) result of thresholding before adjustments, b) final ROI mask.](image)
two touching points between kernels exist, so only two of the lines should be used in the separation. The lines to be considered are the two shortest ones. In Fig. 7c, it is possible to see clearly that the line connecting the two small concave regions would not be appropriate, so only the two others are taken for kernel separation (represented in blue in Fig. 7e). As a result, the pixels in Fig. 7a that correspond to the selected lines are made equal to zero, effectively separating the previously connected kernels (Fig. 7d).

- When four concave regions are present, this almost always indicates the presence of three kernels (under very unlikely conditions, four kernels may be present). In this case, the shortest lines connecting all six possible pairs of regions are determined. The two shortest lines are then selected and used to separate the kernels.

- If five or more regions are present, only the four largest regions are kept, and the procedure described in the previous item is applied. This means that the algorithm is only able to deal properly with clusters with at most three kernels. It is important to highlight that those big clusters will only exist in cases with very high density of kernels, a situation that should be avoided.

2.3.4. Calculation of a Fusarium index
The final step of the algorithm aimed at identifying diseased and sound kernels. Firstly, the best radiation band for the task was chosen by comparing the average pixel values for sound and diseased kernels. The 1411 nm band was chosen as it provided the greatest separation between those average values. After normalisation according to the procedures described previously, this band was used in the calculation of what we define as a Fusarium index (FI), which is given by the proportion of pixels in a kernel with values larger than 0.58. The larger the FI value, the more likely is the presence of FHB. The algorithm was calibrated in such a way an FI value of 0.5 would be the best threshold that separates sound and diseased kernels. However, as it will be seen in Section 3, a relative approach, in which the FI value indicates the likelihood of FHB being present instead of providing a categorical classification, may be more indicated.

2.4. Test setup and validation
Only two images, the first containing only diseased kernels and the second containing only sound kernels, were used to calibrate the algorithm, and the remainder ones were used in the validation tests described next, and whose results are shown in Section 3.

In order to test the performance of the algorithm in detecting FHB in wheat kernels and validate its calibration, the Fusarium index (FI) was calculated for all kernels contained in the 25 images that were not used for calibration of the algorithm. The mean and the standard deviation of the FI obtained for sound and diseased kernels were then calculated. Those values were applied to the Gaussian function expression in order to generate an FI probability density function (PDF) for each group of kernels (sound and diseased). This provides a good visual way of assessing the performance of the algorithm, as the least overlapped are those two distributions, the more successful tends to be the classification. The accuracy of the algorithm, in percentage of correctly classified kernels, was determined by comparing the manually annotated labels with the classification provided by the algorithm, with a kernel being considered diseased if FI > 0.5, and sound otherwise. The results of these tests are shown in Section 3.1.

Tests were also performed using principal component analysis (PCA) as a means of extracting relevant information from all bands in the hyperspectral images. In those tests, the 1411-nm band used in the algorithm would be replaced by one of five images that were generated using the PCA results: the first image was simply the first main component alone, the second one was obtained by averaging the first two main components, the third by averaging the first three main

Fig. 8 shows an example containing several kernel clusters, before and after the application of the cluster split procedure. It can be observed that most separations are very close to ideal, however in a few cases, due to the position of the kernels, small portions of the cluster may be assigned to the wrong kernel. Although undesirable, this error has little effect in the final results, as those misplaced portions are usually less than 10% of the parent kernel areas.
components, and so on. The results achieved by replacing the 1411-nm band with PCA-based images are described in Section 3.1.

In order to assess the ability of the algorithm in estimating DON concentration levels, the FI was calculated for each kernel contained in the six images for which the DON levels were measured in laboratory (see Section 2.2). The correlation between FI values and DON levels was then calculated. The mean and standard deviation of the FI values were calculated for each image and plotted in a graphic showing how those values behave as DON concentrations vary. The objective of this graphic was to determine under which conditions DON concentrations and FI values correlate well, and under which conditions the correlation tends to drop. The results of this analysis are reported in Section 3.2.

Fig. 6 — Separation between kernels and background. a) 672-nm band greyscale representation; b) 1361-nm band greyscale representation; c) 1509-nm band greyscale representation; d) 1657-nm band greyscale representation; e) normalised difference between bands A and D; f) normalised difference between bands B and C; g) combination of the normalised differences; h) mask for the kernels.
As mentioned in the beginning of this section, validation of the algorithm was carried out with the 25 images not used in the calibration of the algorithm. All images used both in the calibration and in the validation were captured under the exact same conditions, so there is not much variability between them. Since images captured under different conditions and setups were not available, it was not possible to perform tests that would further stress the algorithm’s robustness. However, a qualitative analysis on how variations in factors such as illumination and spatial resolution would affect the algorithm could still be performed, with the main remarks being presented in Section 3.1.

3. Results and discussion

As mentioned before, all kernels present in the images were manually annotated prior to being imaged. This manual annotation, which is the reference for all results presented in this section, is subject to errors arising from factors such as fatigue, external distractions and optical illusions, and intra- and inter-rater inconsistencies are common (Nutter, Gleason, Jenco, & Christians, 1993). Thus, this reference is not a ground-truth, a fact that has to be factored when analysing the results.

3.1. FHB detection using the Fusarium index

Figure 9 shows the normal (Gaussian) PDF estimated for sound and diseased kernels. The dots in the plot represent the data summarised in Table 2, that is, the observed distribution of FI values for both types of kernels, considering the 25 images used for validating the algorithm.

Ideally, both PDF curves should be completely disjoint. However, as it can be seen, the curves overlapped moderately for FI values between 0.4 and 0.6, indicating that there is some uncertainty in this range of values. If the threshold value of FI is set at 0.5, that is, a kernel is considered diseased if it has an FI larger than 0.5, and sound otherwise, the accuracy of the algorithm is approximately 91%. A direct comparison with other results in the literature is not possible because the image datasets used in each case are different, and the visual assessments used as basis to calculate the accuracies may be inconsistent. However, a comparison may still be performed if
one takes into consideration that the results are context dependent and not directly extendable to the setups used in the other experiments. Table 3 shows a comparison between the algorithm proposed here and some selected methods found in the literature. As it can be seen, the results achieved by the proposed algorithm are in line with other results reported in the literature.

It is important to emphasise that the labels that are used as reference for the results are subjective and, as such, they are subject to error themselves. Further uncertainty is added by the fact that, when the infection is still in its initial stages, the visual cues of the disease may be nearly imperceptible both to humans and computers. In this context, unless the FI value is very close to zero or one, it is not possible to be 100% certain about a diagnosis; instead, FI indicates a likelihood. It is also important to mention that the uncertainties are larger for individual kernels than for entire batches of seeds: one can be reasonably confident that a kernel is diseased if it has an FI of 0.7, but this conclusion will be almost undeniable if an entire batch of 30 kernels has this same FI value collectively.

Despite the uncertainties surrounding the results, there are several remarks that can be drawn from the tests. The remainder of this section is dedicated to presenting those remarks, emphasising the main aspects that may influence the performance of the algorithm.

Low FI values for diseased kernels are more common when those kernels have a thinner and more elongated shape (Table 4). A possible reason for this is that both cultivars with elongated shape used in this work are moderately resistant to FHB, thus the infection would appear less pronounced in the hyperspectral imaging. Also, because elongated kernels have a greater perimeter to surface area ratio, edge effects such as shadowing and pixel distortion tend to have a greater relative contribution. Another plausible explanation is that this shape makes the visual symptom detection more difficult, harming the visual selection process. Since there is still no hard data to support either possibility, this problem will have to be further investigated for a more conclusive explanation.

High FI values for sound kernels, on the other hand, are more common in kernels with a rounder shape (Table 5). This is majorly because this kind of kernel has a flatter surface, which is more favourable to specular reflection, whose occurrence tends to wash out parts of the image, artificially increasing the value of FI (Fig. 10). A possible solution for this would be taking into consideration only portions of the kernel
that are less subject to specular reflection. However, as stated by Delwiche et al. (2011), taking the whole kernel usually leads to better results.

Shadows cast by the kernels are sometimes taken as part of the kernel instead of being removed with the rest of the background. When this happens, the FI value tends to be lower than it should be. Fortunately, the segmentation procedure adopted in this work is very robust, in such a way only very small portions of the shadow may eventually remain. It is important to mention that shadows could be greatly reduced by using another source of light located opposite to the one used in the experimental setup. However, this solution would increase the problems with specular reflection, and the absence of shadows could potentially make important features in the image to become indiscernible.

According to Delwiche (2003), the orientation of the kernels may influence the results. In this work, this factor did not seem to affect the results. This is probably because the images used for both tuning and testing the algorithm had the respective kernels arranged randomly. If, as suggested by Delwiche (2003), all images had their kernels arranged in a given orientation, the results might be better. However, for practical purposes, it seemed more appropriate to consider the general case in which the user is not concerned with the placement of the kernels.

The procedure to split kernel clusters is effective, but not perfect, as mentioned in Section 2. In order to measure the impact of imperfect divisions on the performance of the algorithm, clustered kernels were separated both by the procedure described in Section 2, and manually. The differences between both approaches were statistically insignificant, which was expected given the relatively mild nature of division imperfections.

As stated before, the algorithm is not capable of dealing with clusters of four or more kernels, so high kernel densities should be avoided, if possible. However, considering that large clusters may eventually occur, the application available for download includes the option of manual cluster separation.

The algorithm uses some rules based on the size of the objects. In principle, those rules are dependent on the spatial resolution of the image, which, in the context of this work, will determine with how many pixels a kernel can be represented. Under this definition, such a resolution will depend both on the pixel resolution of the sensor and on the focal length used in the setup. Considering that most setups will follow certain guidelines, it is expected that the focal length will not vary greatly. Under this assumption, the pixel resolution of the sensor will be the main factor regarding spatial resolution differences. Thus, in order to make the rules based on the size of the objects as robust as possible to setup differences, most of them depend directly on the dimensions (in pixels) of the images. The few rules that are based on absolute values (for example, the 10-pixel threshold for concave regions) were kept this way because they are predicted to work well under any resolution condition expected for this kind of experiment.

Thresholding is an operation that may be very sensitive to lighting conditions. In the algorithm, this operation is applied twice. Since the reflectance values are normalised to lie within the 0–1 range, as described in Section 2.3, most lighting differences that may result from different setups are therefore automatically compensated, so the algorithm should be transferable for different setups without any adjustment. However, it is possible that reflectance relations between kernels, tray and background be different under certain conditions. This should not have a large impact on the first threshold, since the ROI is expected to clearly stand out against a much darker background in any setup. On the other hand, the second threshold is much more dependent on a fine tuning. As a result, if the reflectance relation between kernels and tray are significantly different from that found in this work, a new threshold value may need to be determined.

PCA is often used for extracting relevant information from hyperspectral images. As mentioned in Section 2.4, tests were performed in which the 1411-nm band was replaced by different images generated from the PCA results, with the first main component alone producing the best results. As can be seen in Table 6, the results using the PCA’s first main component were almost identical to those obtained using the 1411-nm band. A probable explanation for this is that all bands generated by the hyperspectral imaging actually have a similar behaviour, that is, diseased kernels tend to appear brighter than sound ones. The main difference among bands is how contrasting diseased kernels appear with respect to background and sound kernels. This contrast variation does not provide a significant amount of additional information that could be captured by PCA and, as a result, the main PCA components tend to converge towards the bands with better contrast, which explains the similar results observed.

As stated earlier, a direct comparison between the proposed algorithm and the methods found in the literature is only possible under the exact same conditions and setup. However, a more loose comparison indicates that the proposed algorithm has an accuracy that is comparable to that achieved by those other methods. Additionally, it presents a set of desirable characteristics that is not commonly found in other methods. Firstly, because it uses only well-known

<p>| Table 5 – Influence of kernel shape on the overestimation of Fusarium, considering only sound kernels. |
|-------------------------------------------------|--------|--------|</p>
<table>
<thead>
<tr>
<th>Sound kernel shape</th>
<th>Mean FI</th>
<th>% Above 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long</td>
<td>0.24</td>
<td>6.5</td>
</tr>
<tr>
<td>Round</td>
<td>0.30</td>
<td>8.0</td>
</tr>
<tr>
<td>All</td>
<td>0.26</td>
<td>7.2</td>
</tr>
</tbody>
</table>

![Fig. 10 — Example of kernel with specular reflection.](image)
morphological operations and very simple ad-hoc rules, it is very simple to implement in any programming language. Some of those languages, like Matlab and Python, already have libraries in which all the required functions are already implemented and available. Secondly, it is computationally light, therefore it may be used even with limited computational power. Thirdly, it can separate clusters of kernels, which is particularly useful when the time available to setup the image capture is limited. Finally, the application is available for download, as explained in the introduction.

3.2. Estimation of DON concentrations

The correlation between FI values and measured DON concentrations was investigated. As stated in Section 2.2, the DON concentrations were calculated for each image as a whole, so an analysis considering each kernel individually was not possible. Figure 11 summarises the results, where the line indicates a hypothetical perfectly linear relationship between DON concentration (see Table 1) and FI (the line appears curved because the abscissa is presented with a logarithmic scale for clarity), the circles indicate the mean FI values for each of the images, and the bars attached to the circles indicate the associated standard deviation.

The measured correlation between FI values and DON concentrations was 84%, which is a significant value. Observing Fig. 11, it can be seen that the correlation is strong for high DON concentrations, and considerably weaker for lower DON levels. This is in line with some findings reported in the literature (Zhou, Kolb, Bai, & Domier, 2002). However, Beyer, Klix, and Verreet (2007) reported that, under certain conditions (e.g. considering higher proportions of damaged kernels), higher correlations can be found. Taking all this into consideration, it seems likely that this mismatch between DON concentrations and visual/FI assessments occurs because, when the DON concentration is low, visible symptoms are either absent or too mild to be visually distinguished. With higher DON concentrations, kernels are more likely to be damaged, thus those variables tend to correlate better.

4. Conclusions

A hyperspectral image processing-based algorithm for the automatic detection of FHB in wheat kernels has been presented. The algorithm is mostly based on widely employed morphological mathematical operations and spectral band manipulations, making it easy to implement and computationally fast. Because the algorithm is designed to be as automated as possible, it does not rely on any human intervention other than submitting the file containing the hyperspectral images. The algorithm is capable of effectively separating clusters composed of up to three kernels. An implementation of the algorithm with a graphical user interface is available for download at https://www.digipathos.cnptia.embrapa.br.

The results of the tests have corroborated the conclusions drawn by other authors that the NIR band is suitable for detecting kernels affected by FHB. Most inconsistencies are due either to specular reflection, or to the inherent uncertainty involved in the manual labelling of the kernels. Because there is no real ground-truth to be pursued, the FI yielded by the algorithm should be used as an indicator of the likelihood

<table>
<thead>
<tr>
<th>Table 6 — Results with the 1411-nm band replaced by the first main component of PCA.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fusarium index (FI) range (% of occurrences)</td>
</tr>
<tr>
<td>0–0.1 0.1–0.2 0.2–0.3 0.3–0.4 0.4–0.5 0.5–0.6 0.6–0.7 0.7–0.8 0.8–0.9 0.9–1.0</td>
</tr>
<tr>
<td>Sound 11.9 25.3 28.0 19.6 8.0 4.3 2.1 0.8 0 0</td>
</tr>
<tr>
<td>Diseased 0 0 1.5 3.1 6.7 7.7 18.1 20.2 28.3 14.4</td>
</tr>
</tbody>
</table>

Fig. 11 — Correlation between DON concentration and FI value.
that a kernel is diseased rather than as a categorical assessment.

Acknowledgements

The authors would like to thank FAPESP (proc. 2013/06884-8) and EMBRAPA (SEG 03.13.00.062.00.00) for funding.

References


