Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Detection of damaged wheat kernels using an impact acoustic signal processing technique based on Gaussian modelling and an improved extreme learning machine algorithm

Min Guo [a,*], *Yuting Ma* [a], *Xiaojing Yang* [a], *Richard W. Mankin* [b]

[a] *Key Laboratory of Modern Teaching Technology, Ministry of Education, School of Computer Science, Shaanxi Normal University, Xi'an, China*

[b] *US Department of Agriculture, Agricultural Research Service, Center for Medical, Agricultural, and Veterinary Entomology, Gainesville, FL, 32608, USA*

ARTICLE INFO

Article history:
Received 2 October 2018
Received in revised form
11 April 2019
Accepted 30 April 2019

Keywords:
Impact acoustics
Extreme learning machine
Damaged wheat kernels
Classification

Wheat kernel damage is a major source of food quality degradation, and long-term feeding on products from damaged wheat kernels will result in malnutrition or even induce diseases. Therefore, detection of damaged wheat kernels is of significant interest. An impact acoustic signal processing technique based on Gaussian modelling and an improved extreme learning machine approach was proposed for detection of insect and sprout-damaged wheat kernels. Discriminant features extracted from Gaussian-model-estimated parameters were fed to an extreme learning machine based on a C-matrix embedded optimisation approximation solution. The best results, 92.0% of undamaged, 96.0% of insect-damaged, and 95.0% of sprout-damaged wheat kernels were correctly classified by using the proposed method. Furthermore, the detection system had good processing speed. Therefore, it could be effective to detect damaged wheat kernels in real time.

1. Introduction

Sprout- and insect-damaged wheat kernels (designated SDK and IDK, respectively) frequently cause degradation of wheat quality. Sprout damage occurs when the moisture at harvest time is too high, causing germination of the seed. This results in sprout shoots emerging from the kernel. Insects inside the wheat kernels also cause nutritional damage to stored grain (Guo, Shang, & Shi, 2005), and contribute excrement and fragments to processed wheat products. The surfaces of the IDK are extremely similar to those of undamaged wheat kernels (UDK), making them difficult to detect by visual observation. Therefore, research on the automatic detection of damaged wheat kernels is of continued urgency.

Although several methods have been studied to detect insects inside wheat kernels, including x-ray imaging, acoustic detection of larval movement and chewing, and carbon

* *Corresponding author.*
E-mail address: guomin@snnu.edu.cn (M. Guo).
https://doi.org/10.1016/j.biosystemseng.2019.04.022

Nomenclature

SDK	sprout damaged wheat kernels
IDK	insect damaged wheat kernels
UDK	undamaged wheat kernels
ELM	extreme learning machine
COAS-ELM	ELM employs a C-matrix embedded optimisation approximation solution
OAS-ELM	ELM employs an optimisation approximation solution ELM
GM	Gaussian model
STFT	short-time Fourier transform
FFT	fast Fourier transform
SLFN	single-hidden layer feedforward network
SVM	support vector machine
BP	back-propagation neural network
RBF	radial basis function
PSO	particle swarm optimisation
CV	cross-validation
GMM	Gaussian mixture model
GISPSA	US Grain Inspection Service, Packers, and Stockyard Administration
t	time-index of frames
T	total number of frames
f	index of the Fourier coefficient
m_X	mean vector
M	covariance matrix of the multidimensional Gaussian distribution
D	dimensionality of the model
σ	covariance diagonal elements
S	training set
$G(x)$	active function
L	number of hidden nodes
w_i, b_i	hidden node parameters
H	hidden layer output matrix
β^+	output weight
I	unit matrix

dioxide measurement, most methods are either labour intensive, expensive, or cannot quantitatively measure insect infestation levels. A method based on impact acoustical signal processing has been studied and has been applied for detection of IDK (Pearson, Cetin, & Tewfik, 2005) since it was successfully used to separate closed pistachio nuts from those with open shells and resulted in approximately 97% classification accuracy (Pearson, 2001). Pearson (2001) used time-domain modelling, short-time variances, and frequency spectra magnitudes as the discriminant features, achieving accuracies of detection of UDK and IDK of 98.0% and 84.4% respectively. To improve the recognition rates, in addition to the methods used before, the maximums in short-time windows and derivative spectra were also adopted as discriminant features, with the result that 98% of UDK and 87% of IDK were correctly detected (Pearson, Cetin, Tewfik, & Haff, 2007).

Subsequently, a new adaptive time-frequency analysis was applied to detect insect-damaged, three types of damaged wheat kernels, reaching 96%, 82%, and 94% accuracies for detection of insect-damaged, pupae-damaged and scab-damaged wheat kernels (Ince et al., 2008). Of course, the method based on impact acoustics was applicable not just for detection of damaged wheat kernels, but for classification of pistachios (Cetin, Pearson, & Tewfik, 2004; Haff & Pearson, 2007; Omid, Mahmoudi, & Omid, 2009; Omid, Mahmoudi, & Omid, 2010; Omid, 2011), hazelnut kernels (Cetin, Pearson, & Sevimli, 2014), walnuts (Khalifa & Komarizadeh, 2012), and rice kernels (Buerano, Zalameda, & Ruiz, 2012).

The prior work indicated that it is feasible to identify damaged wheat kernels by using impact acoustical signal processing, although as stated by Pearson et al. (2007) more study was needed to improve accuracy on kernels infested with insects that have not yet emerged from the kernels. In this paper, a new impact acoustical signal processing technique based on Gaussian modelling and an improved extreme learning machine (ELM) is proposed for detection of damaged wheat kernels. In particular, the study focuses on the problems of detecting IDK and SDK. The main contribution of our work is that the novel method provides high detection accuracies for detection of damaged wheat kernels at rapid speeds.

The improved ELM employs a C-matrix embedded optimisation approximation solution (COAS-ELM) that has better performance than methods based only on an optimisation approximation solution (OAS-ELM) or on traditional ELM. Furthermore, the detection system has low computational cost and rapid processing speed; processing a single wheat kernel requires approximately 30 ms with the throughput rate of approximately 33 kernels s^{-1}, making it feasible to be implemented in a real-time system for sorting damaged wheat kernels.

2. Experimental apparatus

The experimental apparatus for collecting impact acoustic signals of UDK, IDK and SDK wheat kernels (Guo et al., 2016), is shown in Fig. 1, and consists of a vibration feeder, an impact plate, a microphone, and a computer equipped with a sound card. The vibration feeder was used for channelling the bulk wheat kernels into a single-file stream. Freely falling wheat kernels impacted the surface of the impact plate, a stainless steel sheet with dimensions approximately $240 \times 110 \times 0.6$ mm. To avoid wheat kernels bouncing twice before leaving the impact plate, the incline angle was set at 30° above the horizontal through a process of trial and error. The drop distance from the feeder to the impact plate was 500 mm. The impact acoustic signals were collected by an electric condenser microphone (SHURE BG 4.1 with a frequency range of 40 Hz–18 kHz) connected to a computer equipped with a sound card (MAYA44, ESI Audiotechnik GmbH, Leonberg, Germany), which is able to implement the analogue-to-digital conversion of the microphone signals at a sampling frequency of 48,000 Hz. The impact acoustic signals were recorded for later processing.

Hard wheat, whose variety is zhengmai 7698, was used in the tests. The wheat kernels had the same moisture content for all treatments.

Nine hundred visually inspected wheat kernels, 300 each of which were classified as UDK, IDK and SDK, were used for the experiments. Among them, 200 wheat kernels were used for training and 100 were used for testing each class.

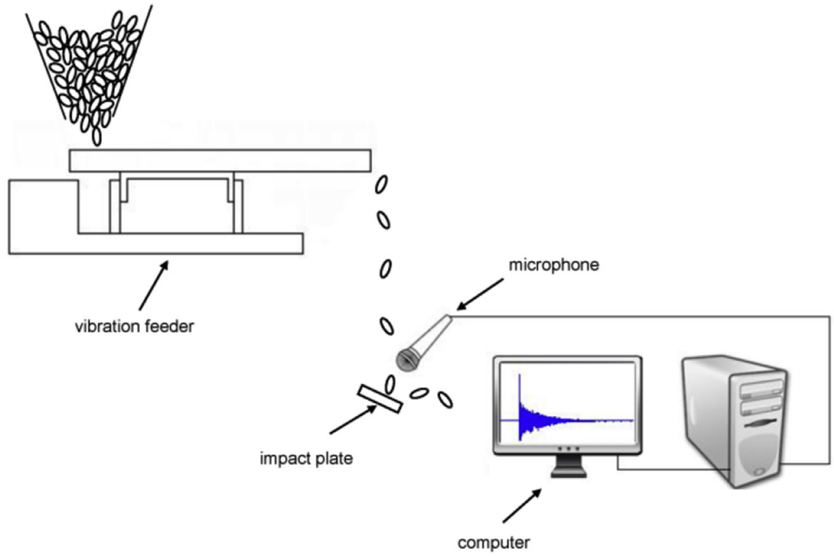

Fig. 1 – The schematic of experimental apparatus.

3. Theory and methods

3.1. *Gaussian model for feature extraction*

Short portions of the impact sound recordings of each kernel can be considered as a stationary or asymptotically stationary random process with Fourier coefficients that are independent complex Gaussian random variables for each frequency bin. This hypothesis was applied for the short-time Fourier transform. A long-term duration of the varying signal was not stationary, but short-duration intervals can be approximated as stationary over one time frame. Consequently Gaussian modelling could be used for each time-frequency bin, then the estimated parameters from the Gaussian model (GM) can be used for feature extraction and further classification. Two steps were conducted for feature extraction: 1) Time-frequency analysis by using the short-time Fourier transform (STFT). 2) Parameters estimation for the GM. The estimated parameters then were used as discriminant features, which were fed to COAS-ELM for classification each of these steps will be demonstrated separately below.

3.1.1. *Time-frequency analysis*

In this paper, 5120 data points were acquired for each impact, beginning 200 points before the maximum magnitude. Typical impact acoustic signals from three types of wheat kernels are shown in Fig. 2 (a), (b), and (c). As is demonstrated, UDK signals may have larger peak values compared with the IDK and SDK, but the maximum amplitudes are so variable they cannot play a decisive role in the detection of each type of wheat kernel. A more obvious characteristic is that the signals from UDK have larger fluctuations during the decaying process, and the decay from IDK and SDK was less stable relative to the signals. This phenomenon may be due to different resonance characteristic of damaged parts of the kernel.

Figure 2 (d), (e), and (f) demonstrate the STFT for each type of wheat kernels. The time-domain signal was segmented in overlapping segments of 512 samples (frames) corresponding to 10.7 ms duration with 50% overlap ratio among subsequent frames. During the calculation of short-time Fourier coefficients, the Hamming window was applied to each frame to reduce the spectral distortions caused by an abrupt change at the boundary points of the segment. Then each windowed frame was fast Fourier transformed by using FFT having the same size of the segment. The STFT processing terminology is as follows: $X(t,f)$ is the complex matrix holding the short-time Fourier transform of each windowed frame, t $(t = 1,\ldots,T)$ is the time-index of frames, where $T = 19$ is the total number of frames, and f $(f = 1,\ldots,F)$ is the index of the Fourier coefficient, where $F = 257$.

3.1.2. *Gaussian modelling*

The GM is widely employed as a statistical model for stationary processes. The probability density function of a GM is defined as:

$$f(x_1,x_2,\ldots,x_n;t_1,t_2,\ldots,t_n) = \frac{1}{(2\pi)^{D/2}|M|^{1/2}} exp\left\{-\frac{1}{2}(X-m_X)^T M^{-1}(X-m_X)\right\} \tag{1}$$

where $X = (x_1,x_2,\ldots,x_n)^T$ is the sample vector, $m_X = (m_X(t_1),m_X(t_2),\ldots,m_X(t_n))^T$ is the mean vector, M is the covariance matrix of the multidimensional Gaussian distribution, and D is the dimensionality of the model.

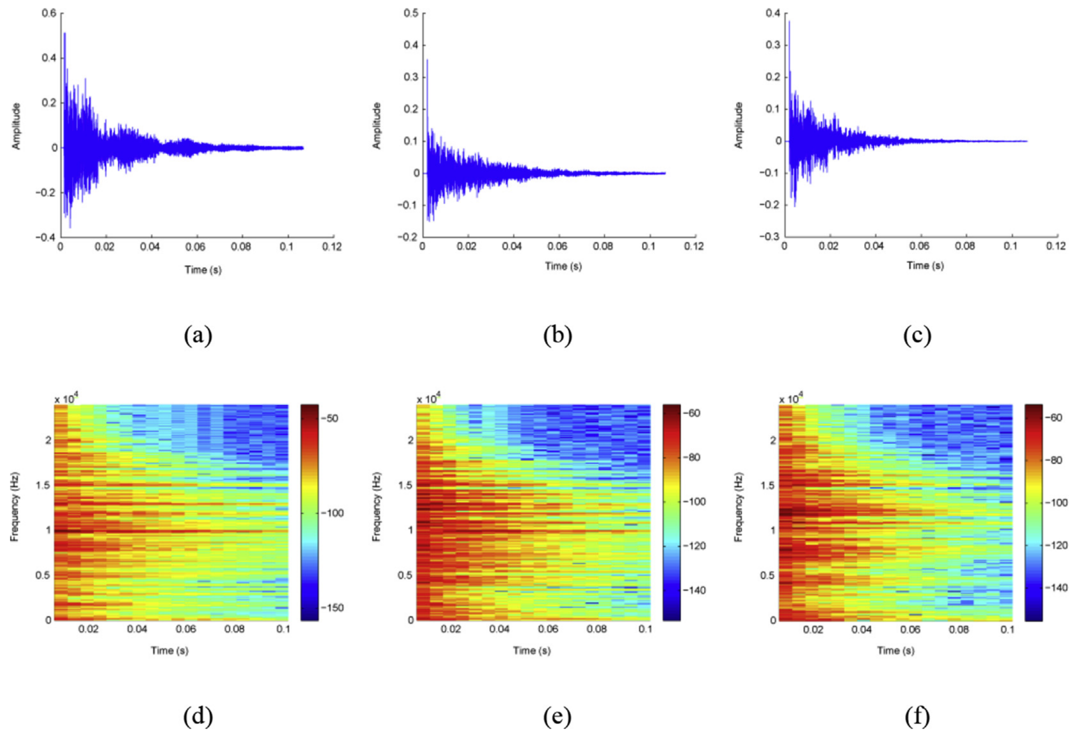

Fig. 2 – Typical impact acoustic signals from (a) UDK, (b) IDK, and (c) SDK, as well as the corresponding time-frequency representations from (d) UDK, (e) IDK (e), and (f) SDK.

It has been shown that diagonal covariance matrix works as effectively as the full covariance matrix, and its use simplifies the analysis. Consequently, a diagonal covariance-based GM was used for parameter estimation, namely

$$M = diag(\sigma_1, ..., \sigma_D) \tag{2}$$

where the σ_1, ..., σ_D are the individual covariance diagonal elements. The parameters of the GM were estimated for each time-frequency bin separately. In total, 38 parameters, 19 mean features and 19 variance features, were used as the discriminant features for further classification.

3.2. COAS-ELM algorithm

3.2.1. ELM based on optimisation approximation solution (OAS-ELM)

In 2006, Huang, Zhu, and Siew (2006) demonstrated that the learning speed of the feedforward neural networks is, in general, far slower than required and has been a major bottleneck in practical applications for past decades. They proposed ELM as a new, faster learning algorithm (Huang et al., 2006). Later, Huang et al. studied ELM for classification in the aspect of the standard optimisation method and extended ELM to a specific type of generalised single-hidden layer feedforward network (SLFN), a support vector network. It was reported that, as analysed in theory and further verified by simulations, classification by ELM tended to achieve better generalisation performance than the traditional support vector machine (SVM) (Huang, Ding, & Zhou, 2010).

As one of the most popular learning algorithms, ELM can achieve good performance in regression and classification. Originally, the solution from ELM algorithm was obtained by solving the Moore-Penrose inverse of hidden layer matrix H. However, in many cases, there is no solution for the equation: $H\beta = T$, Yuan, Wang, and Cao (2011) proposed a new algorithm of ELM based on an optimisation approximation solution that can solve this problem and has better performance than traditional ELM (Yuan et al., 2011). For simplification, it was designated OAS-ELM, and its algorithm can be summarised as follows:

Given a training set $S = \{(x_1, t_1), ..., (x_N, t_N)\}$, active function $G(x)$, and the number of hidden nodes L; $\beta = (\beta_1, \beta_2, ..., \beta_L)^T, Y = (t_1, t_2, ..., t_N)^T$;

Step 1. Randomly assign hidden node parameters $\omega_i, b_i (i = 1, 2, ..., L)$;

Step 2. Calculate the hidden layer output matrix H, let $r = rank(H)$;

Step 3. if $r = L$, calculate the output weight:

$$\beta^+ = (H^T H)^{-1} H^T T \tag{3}$$

if $r = N$, calculate the output weight:

$$\beta^+ = H^T (HH^T)^{-1} T \tag{4}$$

Otherwise, let $K = H^T H, c = H^T T, \varepsilon = 10^{-4}$, solve the optimal models $\min||K - B||$, $\min||B^* \beta - c||$, and obtain the optimal solutions B^*, β^*;

If $||c - K\beta^*|| > \varepsilon$ and $||c - B^*\beta^*|| > \varepsilon$, calculate

$$B = B^* + \frac{(c - B^*\beta^*)(c - B^*\beta^*)^T}{(c - B^*\beta^*)^T\beta^*} \tag{5}$$

Else calculate

$$B = B^* + \frac{cc^T}{c^T\beta^*} - \frac{B^*\beta^*(\beta^*)^T B^*}{(\beta^*)^T B^*\beta^*} \tag{6}$$

Then calculate the output weight $\beta^+ = B^{-1}c$.

3.2.2. A further improved ELM (COAS-ELM)

Considering the standard model for multiple linear regression:

$$T = H\beta + \varepsilon \tag{7}$$

Using unbiased linear estimation with minimum variance or maximum likelihood estimation when the random vector, ε, is normal, gives

$$\widehat{\beta} = (H^T H)^{-1} H^T T \tag{8}$$

as an estimate of β. The minimum sum of squares of the residuals is:

$$\varphi(\beta) = (T - X\widehat{\beta})^T (T - X\widehat{\beta}) \tag{9}$$

Unfortunately, the least squares estimate in Eq. (9) suffers from the deficiency of mathematical optimisation techniques that give point estimates, the estimation procedure does not include a method for portraying the sensitivity of the solution Eq. (9) to the optimisation criterion Eq. (10) (Hoerl & Kennard, 1970).

Hoerl (1962) first suggested that to control the inflation and general instability associated with the least squares regression estimates, one can use

$$\widehat{\beta}* = (H^T H + cI)^{-1} H^T T = (H^T H + C)H + C)^{-1} H^T T \tag{10}$$

where I denotes the unit matrix, $C = cI$, and $c > 0$. The family of estimates given by $c > 0$ has many mathematical similarities with the portrayal of quadratic response functions. For this reason, estimation and analysis built around Eq. (10) has been labelled "ridge regression" (Hoerl, 1962).

Estimation based on the matrix $[H^T H + cI]$ rather than on $H^T H$ has been found to be a procedure that can be used to help circumvent many of the difficulties associated with the usual least squares estimates, in particular, the procedure can be used to portray the sensitivity of the estimates to the particular set of data being used, and it can be used to obtain a point estimate with a smaller mean square error (Hoerl & Kennard, 1970).

Therefore, in this paper, Eqs. (3) and (4) can be improved as follows:

If $r = L$, calculate the output weight:

$$\beta^+ = (H^T H + cI)^{-1} H^T T \tag{11}$$

if $r = N$, calculate the output weight:

$$\beta^+ = H^T (HH^T + cI)^{-1} T \tag{12}$$

As in previous analysis, $C = cI$ is added to HH^T or $H^T H$ to obtain a smaller mean square error and better generalization performance (Hoerl & Kennard, 1970). The new ELM, based on the C-matrix embedded optimisation approximation solution, is named as COAS-ELM.

4. Results

4.1. COAS-ELM for classification

Thirty eight parameters extracted by Gaussian modelling were used as the discriminant features that were fed to COAS-ELM subsequently for classification. To obtain better generalisation performance, the value of c in Eqs. (11) and (12) was set to 390, the number of hidden nodes of COAS-ELM was designated as 140. The sigmoid function was used as the activation function. Further details of parameter selection are presented in the discussion.

Table 1 shows the classification results for the three types of wheat kernels using the proposed method. Under the best results, 92.0% of UDK, 96.0% of IDK and 95.0% of SDK were correctly classified. The results indicated that it was effective to detect damaged wheat kernels by using the proposed method, the good results not only attribute to the feasibility of feature extraction but also the good generalization performance of COAS-ELM.

4.2. Comparison of ELM, OAS-ELM and COAS-ELM

A comparative study of ELM, OAS-ELM and COAS-ELM was conducted, which used 50 separate trials using with the same 900 kernels. For each trial, the learning time, the average recognition rates both for training and testing were computed. Figure 3 shows the results, which indicate that COAS-ELM can achieve higher testing accuracies, and there was little difference among learning times for the different types of ELM, which was attributed, primarily, to the rapid learning speed of ELM. Table 2 presents the statistics from the 50 trials.

4.3. Comparison of various classification methods

Further study was conducted to compare various classification methods. As Table 3 shows, the classifier based on COAS-ELM reached better average detection accuracy more rapidly than the other methods. For the back-propagation (BP) neural network, a 38-8-3 BP configuration was adopted which

Table 1 – Classification results for three types of wheat kernels by using the proposed method.

Damage category	Classification outcomes (%)		
	Undamaged	IDK	Sprouted
Undamaged	92.0	6.0	2.0
IDK	1.0	96.0	3.0
Sprouted	5.0	0.0	95.0

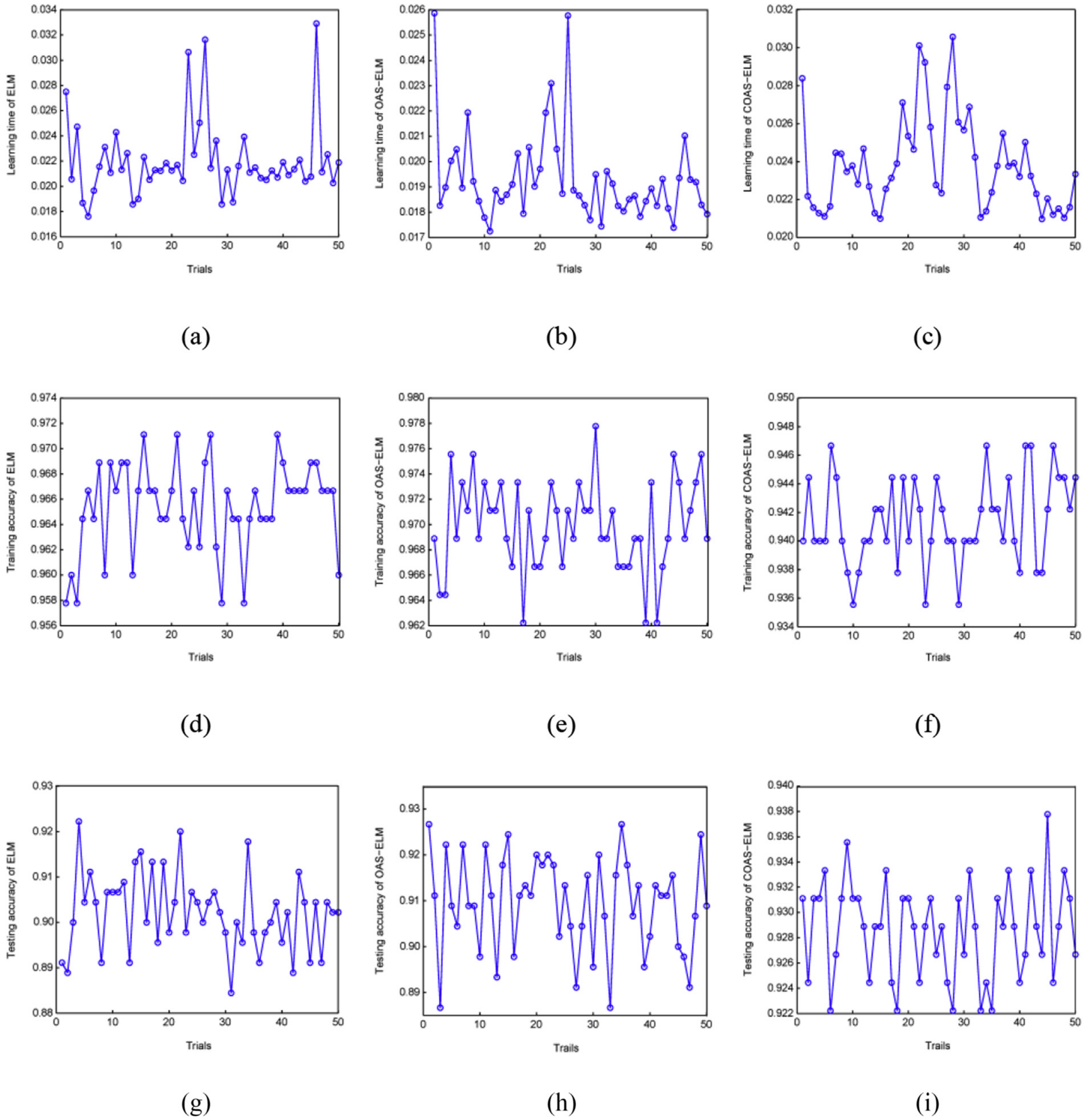

Fig. 3 – Performance comparisons among ELM, OAS-ELM, and CLAS-ELM for different trials of 900 kernels.

exhibited highest recognition rates through trial and error, that is, BP had 38 neurons in input layer, 8 neurons in hidden layer, and 3 neurons in output layer.

For the support vector machine classifier, the SVM software of Libsvm3.23 (https://www.csie.ntu.edu.tw/~cjlin/) was used for multi-classification, and the radial basis function (RBF) was adopted as the kernel function. As Table 3 shows, when we used the default values of the kernel parameters, detection accuracies for UDK, IDK, and SDK were 93.0%, 92.0%, and 88.0% respectively, and the average accuracy was just 91.0%. When particle swarm optimisation (PSO) was used for selecting the optimal penalty factor and kernel parameter,

Table 2 – Performance statistics for three types of ELM.

ELM Type	Average training accuracy	Average testing accuracy	Average learning time (s)
Traditional ELM	0.9654	0.9023	0.0220
OAS-ELM	0.9699	0.9095	0.0193
COAS-ELM	0.9414	0.9286	0.0238

Table 3 – Results for various classification methods.

Methods	Accuracies (%)				Running time (s)
	Undamaged	IDK	Sprouted	Average	
Method in this paper	92.0	96.0	95.0	94.3	0.032
STFT-GM-ELM	92.0	96.0	90.0	92.7	0.029
STFT-GM-BP	91.0	96.0	88.0	91.7	0.500
STFT-GM-SVM	93.0	92.0	88.0	91.0	0.053
STFT-GM-PSO-SVM	96.0	94.0	91.0	93.7	889.29

finding the one giving the highest ten-folds cross-validation (CV) accuracy of the training set, 96.0% of UDK, 94.0% of IDK, and 91.0% of SDK were correctly detected, and the average accuracy was 93.7%, which was nearly as good as the proposed method. However, the PSO algorithm was very time-consuming. When 600 wheat kernels with 200 of each type were used for training, the procedure took approximately 889s. Detection of damaged wheat kernels requires high throughput rates, and a classifier based on PSO-SVM could not be used in real-time, so this method was considered unacceptable.

5. Discussion

It should be noticed that in this paper GM was used rather than a Gaussian mixture model (GMM) for parameter estimation. Although a GMM may be regarded as a better parameter approximation scheme, this increases the amount of computation of the system, resulting in lower processing speeds. For the GMM, the direct estimation of the parameters by maximum likelihood method is EM-algorithm, which suffers from slow convergence. Park and Ozeki (2009) analysed the dynamics of the EM algorithm for Gaussian mixtures around singularities and showed that there exists a slow manifold caused by a singular structure, which is closely related to the slow convergence of the EM algorithm. For comparison, each time-frequency bin was adapted to fit GMMs with just two components following an expectation maximisation procedure. The processing time for the GMM parameter estimations for each time-frequency bin was approximately 0.32s. Thus, at most only 3 kernels can be classified and recognised per second. However, detection of damaged wheat kernels requires high throughput, therefore this processing speed is unacceptable and is difficult to apply. When GM was applied, processing the impact from each wheat kernel, including feature extraction and classification, required approximately 30 ms. This corresponds to a processing speed of approximately 33 wheat kernels per second, making it feasible for application. For example, The US Grain Inspection Service, Packers, and Stockyard Administration (GISPSA) guidelines classify samples through sieving and visually inspecting a sample (1 kg) to detect the presence of insects and determine the quality of a particular shipment, which are slow and labour intensive. Thus the method proposed in this paper for the detection of damaged wheat kernels is suitable for this work.

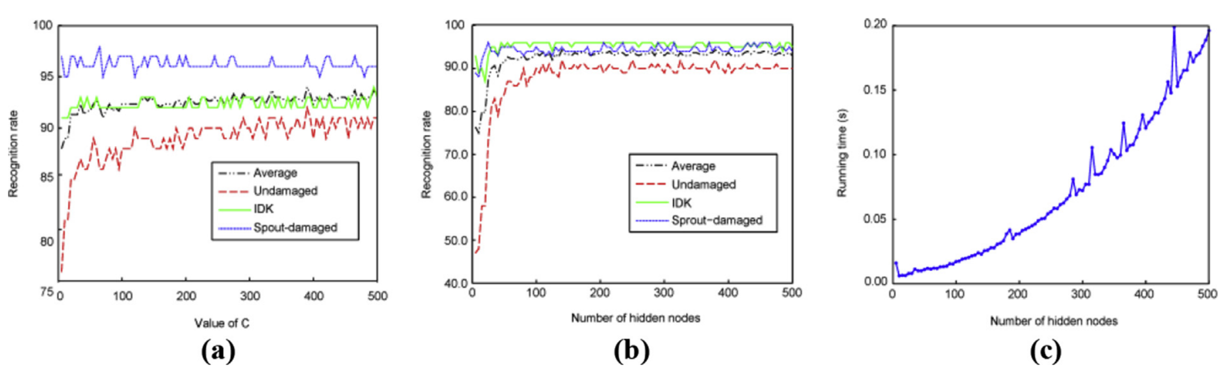

Fig. 4 – COAS-ELM parameters selection.

Table 4 – Accuracies for various activation functions.

Activation functions	Accuracies (%)				Running time (s)
	Undamaged	IDK	Sprouted	Average	
Sigmoid	92.0	96.0	95.0	94.3	0.032
Sin	92.0	92.0	93.0	92.3	0.031
Hardlim	73.0	75.0	84.0	77.3	0.032
Tribas	85.0	86.0	92.0	87.7	0.029
Radbas	87.0	91.0	95.0	91.0	0.033

In addition, the value of c in Eqs. (7) and (8) is important and contributes to improved generalisation performance. In Fig. 4 (a), values of c that vary from 5 to 500 in steps of 5 were tested in order to acquire the optimal classification performance. According to the figure, the higher average recognition rate mainly exists in the regions where the values of c are large. When the values of c increase, recognition rates remain stable, and the average recognition rate reaches its highest value when c is 390. Thus, the value of c was set to 390 in the experiment. Also, the number of hidden nodes should be taken into consideration. As the number of hidden nodes increases in Fig. 4(b), the accuracies for the three types also increase, especially for SDK. When the number of hidden nodes was large, the accuracies for three types were satisfying and basically remained stable. Zhu, Qin, Suganthan, and Huang (2005) showed that ELM may need higher numbers of hidden neurons due to the random determination of the hidden node parameters and hidden biases (Zhu et al., 2005). Because the inclusion of hidden nodes increases running time (Fig. 4 (c)), 140 hidden nodes were adopted in the experiment. Furthermore, the influence of various activation functions was considered as Table 4 shows, and ultimately the sigmoid function was used as the activation function.

6. Conclusion and future work

A novel impact acoustic processing technique based on Gaussian modelling and an improved ELM, based on a C-matrix embedded optimisation approximation solution (COAS-ELM) was proposed for detection of damaged wheat kernels. Fourier analysis was performed on the impact signals and discriminant features were extracted from the parameters estimation of a GM for each time-frequency bin. The discriminant features were fed to the COAS-ELM and damaged kernels thereby were discriminated from undamaged kernels. The best results, 92.0% of UDK, 96.0% of IDK and 95.0% of SDK were correctly classified, which indicated the effectiveness of the proposed method. Furthermore, the detection system had high processing speed. Processing a single wheat kernel requires approximately 30 ms, with the throughput rate of approximately 33 kernels s^{-1}. This method is therefore easily trainable and thus can be the basis of the future research. For example, detection performance can be evaluated in additional circumstances, such as different temperatures, different sizes of wheat kernels, and different levels of maturity.

Acknowledgments

This work was supported by the Science Research and Development Program of Shaanxi Province of China (No. 2016NY-176).

REFERENCES

Buerano, J., Zalameda, J., & Ruiz, R. S. (2012). Microphone system optimization for free fall impact acoustic method in detection of rice kernel damage. *Computers and Electronics in Agriculture, 85*(5), 140–148.

Cetin, A. E., Pearson, T. C., & Sevimli, R. A. (2014). System for removing shell pieces from hazelnut kernels using impact vibration analysis. *Computers and Electronics in Agriculture, 101*(2), 11–16.

Cetin, A. E., Pearson, T. C., & Tewfik, A. H. (2004). Classification of closed-and open-shell pistachio nuts using voice-recognition technology. *Transactions of the ASAE, 47*(2), 659–664.

Guo, M., Ma, Y., Zhao, Z., Ma, M., Wu, X., & Mankin, R. W. (2016). A new EEMD-based scheme for detection of insect damaged wheat kernels using impact acoustics. *Acta Acustica united with Acustica, 102*(6), 1108–1117.

Guo, M., Shang, Z., & Shi, H. (2005). Sound absorption measurements of various types of grain. *Acta Acustica united with Acustica, 91*(5), 915–919.

Haff, R. P., & Pearson, T. C. (2007). Separating in-shell pistachio nuts from kernels using impact vibration analysis. *Sensing and Instrumentation for Food Quality and Safety, 1*(4), 188–192.

Hoerl, A. E. (1962). Application of ridge analysis to regression problems. *Chemical Engineering Progress, 58*(3), 54–59.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67.

Huang, G. B., Ding, X., & Zhou, H. (2010). Optimization method based extreme learning machine for classification. *Neurocomputing, 74*(1–3), 155–163.

Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing, 70*(1), 489–501.

Ince, N. F., Onaran, I., Pearson, T. C., Tewfik, A. H., Cetin, A. E., Kalkan, H., et al. (2008). Identification of damaged wheat kernels and cracked-shell hazelnuts with impact acoustics time-frequency patterns. *Transactions of the ASABE, 51*(4), 1461–1469.

Khalifa, S., & Komarizadeh, M. H. (2012). An intelligent approach based on adaptive neuro-fuzzy inference systems (ANFIS) for walnut sorting. *Australian Journal of Crop Science, 6*(2), 183–187.

Omid, M. (2011). Design of an expert system for sorting pistachio nuts through decision tree and fuzzy logic classifier. *Expert Systems with Applications, 38*(4), 4339–4347.

Omid, M., Mahmoudi, A., & Omid, M. H. (2009). An intelligent system for sorting pistachio nut varieties. *Expert Systems with Applications, 36*(9), 11528–11535.

Omid, M., Mahmoudi, A., & Omid, M. H. (2010). Development of pistachio sorting system using principal component analysis (pca) assisted artificial neural network (ann) of impact acoustics. *Expert Systems with Applications, 37*(10), 7205–7212.

Park, H., & Ozeki, T. (2009). Singularity and slow convergence of the EM algorithm for Gaussian mixtures. *Neural Processing Letters, 29*(1), 45–59.

Pearson, T. C. (2001). Detection of pistachio nuts with closed-shells using impact acoustics. *Applied Engineering in Agriculture, 17*(2), 249–253.

Pearson, T. C., Cetin, A. E., & Tewfik, A. H. (2005). Detection of insect damaged wheat kernels by impact acoustics. In *, Vol. 5. Proceedings (ICASSP 05). IEEE international conference on acoustics, speech, and signal processing* (pp. 649–652).

Pearson, T. C., Cetin, A. E., Tewfik, A. H., & Haff, R. P. (2007). Feasibility of impact-acoustic emissions for detection of damaged wheat kernels. *Digital Signal Processing, 17*(3), 617–633.

Yuan, Y., Wang, Y., & Cao, F. (2011). Optimization approximation solution for regression problem based on extreme learning machine. *Neurocomputing, 74*(16), 2475–2482.

Zhu, Q. Y., Qin, A. K., Suganthan, P. N., & Huang, G. B. (2005). Evolutionary extreme learning machine. *Pattern Recognition, 38*(10), 1759–1763.