



## OPEN ACCESS

## EDITED BY

Yanbei Zhu,  
National Institute of Advanced  
Industrial Science and Technology  
(AIST), Japan

## REVIEWED BY

Zhiming Guo,  
Jiangsu University, China  
Leiqing Pan,  
Nanjing Agricultural University, China

## \*CORRESPONDENCE

Liping Chen  
chenlp@nercita.org.cn  
Qingyan Wang  
wangqy@nercita.org.cn

## SPECIALTY SECTION

This article was submitted to  
Technical Advances in Plant Science,  
a section of the journal  
Frontiers in Plant Science

RECEIVED 31 May 2022

ACCEPTED 28 July 2022

PUBLISHED 15 September 2022

## CITATION

Zhang C, Huang W, Liang X, He X,  
Tian X, Chen L and Wang Q (2022)  
Slight crack identification of  
cottonseed using air-coupled  
ultrasound with sound to image  
encoding. *Front. Plant Sci.* 13:956636.  
doi: 10.3389/fpls.2022.956636

## COPYRIGHT

© 2022 Zhang, Huang, Liang, He, Tian,  
Chen and Wang. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Slight crack identification of cottonseed using air-coupled ultrasound with sound to image encoding

Chi Zhang<sup>1</sup>, Wenqian Huang<sup>1</sup>, Xiaoting Liang<sup>1,2</sup>, Xin He<sup>1</sup>,  
Xi Tian<sup>1</sup>, Liping Chen<sup>1\*</sup> and Qingyan Wang<sup>1\*</sup>

<sup>1</sup>Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, <sup>2</sup>College of Information Technology, Shanghai Ocean University, Shanghai, China

Slight crack of cottonseed is a critical factor influencing the germination rate of cotton due to foamed acid or water entering cottonseed through testa. However, it is very difficult to detect cottonseed with slight crack using common non-destructive detection methods, such as machine vision, optical spectroscopy, and thermal imaging, because slight crack has little effect on morphology, chemical substances or temperature. By contrast, the acoustic method shows a sensitivity to fine structure defects and demonstrates potential application in seed detection. This paper presents a novel method to detect slightly cracked cottonseed using air-coupled ultrasound with a light-weight vision transformer (ViT) and a sound-to-image encoding method. The echo signal of air-coupled ultrasound from cottonseed is obtained by non-contact and non-destructive methods. The intrinsic mode functions (IMFs) of ultrasound signal are obtained as the sound features using variational mode decomposition (VMD) approach. Then the sound features are converted into colorful images by a color encoding method. This method uses different colored lines to represent the changes of different values of IMFs according to the specified encoding period. A light-weight MobileViT method is utilized to identify the slightly cracked cottonseeds using encoding colorful images corresponding to cottonseeds. The experimental results show an average overall recognition accuracy of 90.7% for slightly cracked cottonseed from normal cottonseed, which indicates that the proposed method is reliable to applications in detection task of cottonseed with slight crack.

## KEYWORDS

crack cottonseed identification, variational mode decomposition, sound to image encoding, vision transformer, deep learning, air-coupled ultrasound

## Introduction

Cotton is an important economic crop throughout the world. The quality of cottonseed is an important factor in determining the yield and quality of cotton. Cotton seed will go through a series of processes such as ginning and stripping, which will cause a lot of damage to cotton seed. However, in the process of removing excess linters of

cottonseed, the foamed acid will enter the cottonseed through the cracks and diminish the germination of the cotton seed. After sowing, water will also enter cottonseed through cracks, further reducing the germination of cottonseed. Therefore, cracked cottonseeds will decrease cotton yield.

To reduce the amount of cottonseed wasted, the automation system of cotton precision seeding is often applied in practical production. After precision seeding, there is no need for thinning seedlings or avoiding inconsistency of individual growth and development in cotton field. Therefore, cotton precision seeding can significantly decrease the production cost of cotton, improve the efficiency of field management, and consequently, realize standardized planting. However, this technology puts forward higher requirements for the quality of cottonseeds, which makes the quality detection of cottonseeds crucial. The traditional seed detection method is destructive, inefficient, time-consuming, and non-automated. Developing fast and high-throughput non-destructive detection methods for seed quality is urgently needed for agricultural production. In recent years, non-destructive detection technologies, such as machine vision, optical spectroscopy, thermal imaging, and acoustics, have gradually become new ways to detect seed quality.

Machine vision is a rapid and non-destructive technology and has been applied to detect the quality and safety of seed. Based on morphological and color features extracted from images acquired by camera with high resolution, morphology (Rodríguez-Pulido et al., 2012), color (Tu et al., 2018), shape (Li et al., 2016), size, texture, and exterior defects (Huang et al., 2019) of seed can be evaluated. Severely damaged and broken cottonseeds can also be identified effectively by extracting morphological characteristics from image (Bai et al., 2018). However, injuries recognition and location are still major difficulties in using machine vision for seed surface defect detection (Huang et al., 2015). Slight injuries such as small cracks in cottonseed are hardly distinguished by vision. Moreover, it is particularly difficult for imaging the injuries on the edge and the back side of seed that are hidden from the camera's field of view.

Optical spectroscopy is also a powerful tool to inspect seed, especially in characterizing internal quality. Based on the interactions between light and material molecular groups, the information of corresponding chemical compositions can be extracted from the changes of optical spectra. According to the form of interaction, spectroscopic techniques are mainly based on light absorption (e.g., near-infrared spectroscopy), light scattering (e.g., Raman spectroscopy) and light emission (e.g., fluorescence spectroscopy). Most of them have been used to determine seed quality and safety, including internal compositions (Sunoj et al., 2016), moisture (Zhang and Guo, 2020), germination (Fan Y. et al., 2020) and infection (Tao et al., 2019). Hyperspectral imaging technique incorporates optical spectroscopy and imaging technology to obtain spatial and spectroscopic information simultaneously. Combined

with machine learning methods, such as linear discriminant analysis (LDA), partial least-squares discriminant analysis (PLS-DA), support vector machine (SVM), and artificial neural networks (ANN), a hyperspectral imaging data cube can provide the information of chemical compositions and their distributions, which makes this a potential technology in the seed industry, especially in variety identification (Zhou et al., 2020), classification (Barboza da Silva et al., 2021) and chemical composition determination (Yang et al., 2018; Hu et al., 2021). Spectroscopic technology also does a good job in damage detection of seed, such as insect damage (Chelladurai et al., 2014), fungi damage (Baek et al., 2019), frost damage, and sprout damage, because all these types of damage can lead to the change of chemical compositions of seed, which could change the spectral features. Nevertheless, it is still challenging to distinguish pure physical damage with little chemical change, such as slight crack in cottonseed.

Thermal imaging is a non-destructive technique for converting the invisible infrared radiation pattern of an object into visible images for feature extraction and analysis (Rahman and Cho, 2016). Based on the changes of surface temperature, the infrared radiation profile of seed can be mapped and analyzed. Unlike above-mentioned methods, no illumination sources are required in this system, and a thermal imaging camera along with its data acquisition system is enough to provide information of object. Thermal imaging technology has found its way in estimating seed quality, including determination of morphological features, detection of diseases and insect infestation, evaluation of viability (Belin et al., 2018) and germination performance (Fang et al., 2016), distinguishing aged or dead seeds from healthy ones (Kim et al., 2014), and monitoring seed quality during storage (Xia et al., 2019). In addition, thermal imaging has the capability of sensing all possible physical damage of seed, since there is a significant relationship between seed temperature and degree of damage (ElMasry et al., 2020). But the difficulties of detecting slight physical injuries, which are hard to recognize by machine vision and even human vision, still exist for the thermal imaging method.

An acoustic method is also developed for non-destructive detection of agricultural products. Among acoustic technology, ultrasonic testing is an important method. With the advantages of short wavelength, high frequency, and good directional property, ultrasound possesses better penetrability than audible sound and subsonic wave, and consequently becomes a powerful tool in non-destructive testing of seed. Ultrasound signal produced by impacting seed can be used to evaluate seed quality. When crack appears on seed testa, the structural strength and damping coefficient of the seed will change, which leads to the variations of frequency and intensity of the impacting ultrasound signal. Depending on the differences in echo signals between healthy seed and defective seed, the acoustic method shows more superiority in recognizing fine surface crack than

common non-destructive detection methods. The approach was first proposed to distinguish pistachio nuts with open shells from those with closed shells (Pearson, 2001), and the results showed that the detection accuracy of pistachio nuts was ~97%. Combined with signal processing and identifying algorithm, the acoustic method is applied to detect insect damage (Pearson et al., 2007; Yanyun et al., 2016) and mildew damage (Sun et al., 2018) of seed. The potential to identify fine defects with this method is expected and verified. However, detecting light crack of cottonseed with smaller size than most seeds is rarely reported.

In this paper, a non-destructive detection method based on an air-coupled ultrasonic inspection system is developed to distinguish cottonseed with slight crack from intact kernel. VMD is utilized to decompose an ultrasonic signal into band-limited multiple IMFs. These IMFs are used to construct the feature matrix of ultrasonic signal. Then the feature matrix is converted into the colorful image using a color encoding method. A deep learning-based method combining Transformer model with CNN model is used to classify the color images generated from air-coupled ultrasonic cottonseeds. Finally, the performance of the proposed method is compared with other detection methods.

## Materials and methods

### Samples

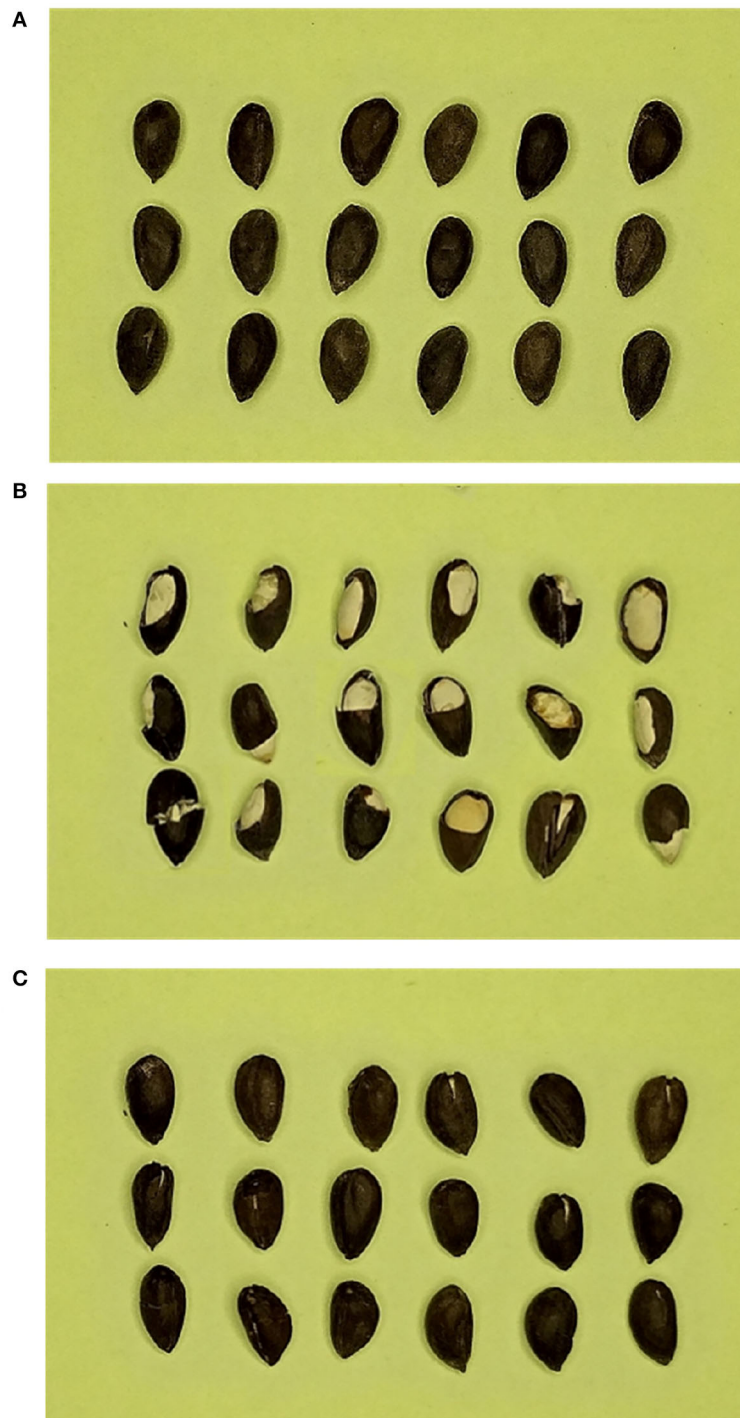
A total of 600 cottonseeds named Xinluzhong52 are used in this study. The total cottonseeds consisted of 296 intact kernels and 304 kernels with slight crack. Cracked cottonseed is damaged cottonseed showing the obvious white endosperm inside. For slightly cracked cottonseed, it is difficult to detect the endosperm inside, but there is crack on testa of cottonseed. Physical images of cottonseed are shown in Figure 1. The cottonseed with severe crack that shows the white endosperm can be identified by machine vision method. This research focuses on the method of distinguishing cottonseed with slight crack from intact cottonseed. Therefore, the samples only include the two classes of cottonseeds.

### Detection system based on air-coupled ultrasound

To maintain sufficient energy transmission, liquid couplant, such as water, oil etc. is used to immerse samples in a traditional ultrasonic technique. The mode of contact coupling may cause damage or pollution on the surface of sample. The air-coupled ultrasonic technique emerges as a novel approach

for non-destructive, non-contact, and rapid inspection (Fang et al., 2017). The surrounding air is used as the couplant between transmitting transducer and materials or between receiving transducer and materials in air-coupled ultrasonic techniques. The significant advantage of an air-coupled ultrasonic technique is avoiding the use of traditional couplant. Therefore, it becomes a reliable and effective non-destructive detection method. The air-coupled ultrasonic technique is suitable for industrial detection applications, such as the natural defects in wood (Tiitta et al., 2020), corn seed with hole (Yanyun et al., 2016) and food engineering (Fariñas et al., 2021a,b).

The air-coupled ultrasonic detection system is used to obtain the ultrasonic echo signal. The inspection system is shown in Figure 2. The signal acquisition system consists of a pair of transducers (400K-20N-R50-T and 400K-20N-R50-R, PR, China), a preamplifier (400K, PR, China), an air-coupled ultrasonic inspection instrument (PRACUT-111, PR, China), and an industrial computer. The center frequency of the transducer is 400 kHz. The diameter of the piezoelectric ceramic disc is 20 mm. The focal length of the transducer is 50 mm. The normal through-transmission mode is applied to two transducers. In order to obtain accurate ultrasonic signals, the two transducers need to be strictly aligned. The cottonseed is placed as the focus between transmitting transducer and receiving transducer. To meet the requirement of cottonseed placement, an aluminum plate with a thickness of 0.2 mm as a holder is fixed in the middle of two transducers. The thin aluminum plate can guarantee that the air-coupled ultrasonic signal transmitted by the transmitting transducer can be received by the receiving transducer as much as possible. The preamplifier with the amplification factor of 60 dB is used to amplify and filter the ultrasonic signal received by the air-coupling receiver transducer (400K-20N-R50-R), and then input the processed ultrasonic signal to the air-coupled ultrasonic inspection instrument. The ultrasonic inspection instrument outputs 400 V excitation signal to drive the transmitting transducer and converts the received ultrasonic signal into digital signal. The industrial computer mainly consists of Intel i7-6700 CPU @3.4 GHz, 512G SSD hard drive, 16 GB RAM and a GPU (Geforce RTX 2080 WindForce OC 8G, GIGABYTE). The computer sends the control command to the ultrasonic inspection instrument through the USB interface, and receives the ultrasonic data sent by the ultrasonic inspection instrument through a LAN interface. The ultrasonic data are stored in the computer using PRACUT software (Suzhou Phaserise Technology Co., Ltd., China). Although Python language and Pytorch deep learning framework are used to realize the signal processing and the identification of cottonseed with slight crack, real-time detection can be realized by C++ language with the standard dynamic link library provided by PRACUT software.

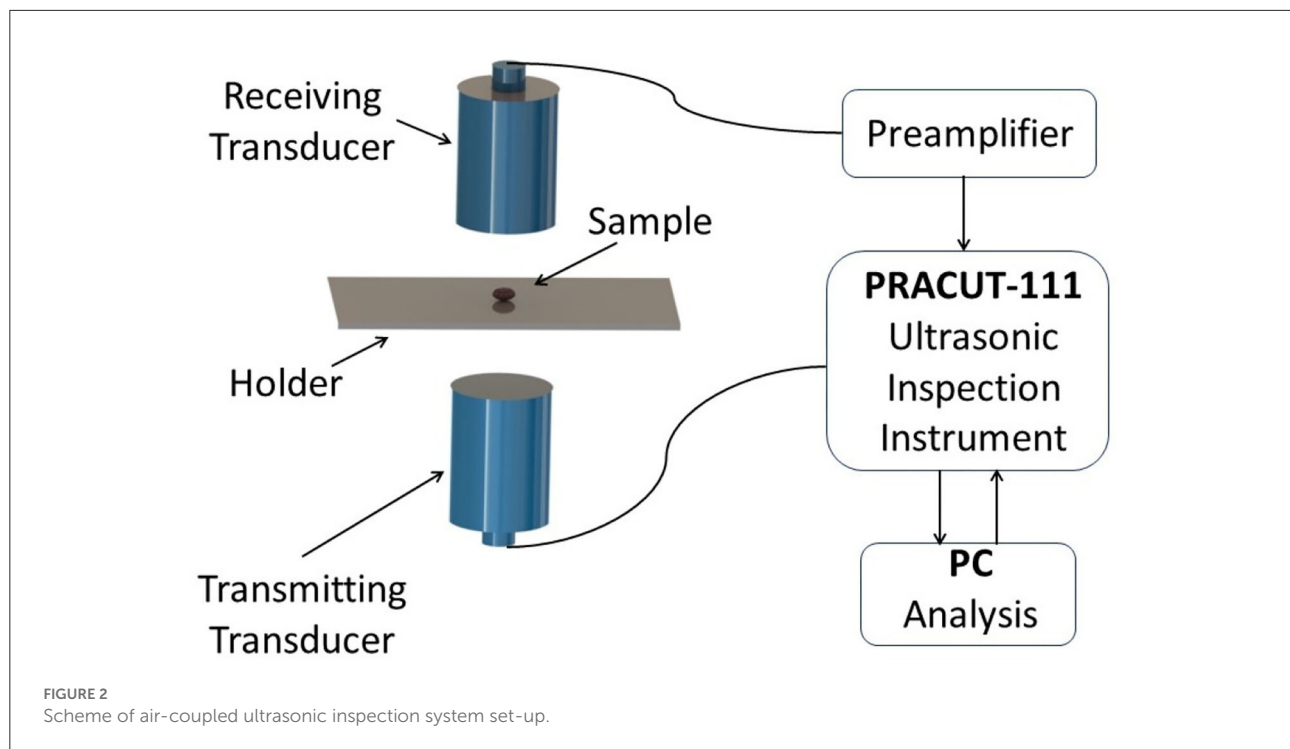


**FIGURE 1**  
Cottonseed kernels. (A) Intact cottonseeds. (B) Cottonseeds with severe cracks. (C) Cottonseeds with slight cracks.

## Ultrasonic signal acquisition

A-scan mode of the air-coupled ultrasonic inspection instrument is used to identify cottonseed with slight crack.

In A-scan mode, damaged cottonseeds are detected through changes in air-coupled ultrasonic signal passing through them. In order to generate the ultrasonic signal data set, first, the cottonseed sample is placed on the aluminum plate so that it



coincides with the focus position of the transducers. Then the ultrasonic signal is obtained after the ultrasonic signal passes through the cottonseed. The ultrasonic signal data is exported and saved as a CSV format file. Finally, the category label (1 or 0) corresponding to each sample is appended to the end of CSV file, where “1” represents the normal cottonseed and “0” represents the cottonseed with slight crack. Each cottonseed sample corresponds to a CSV file, and all CSV files constitute the data set used in this study. The typical ultrasonic signals from intact cottonseed and cottonseed with slight crack are shown in Figure 3 respectively. The amplitude fluctuations of the two types of ultrasonic signals are very similar, so it is very important to extract effective features for classification from these signals.

## Identification of slight crack cottonseed

### Variational mode decomposition method

Variational mode decomposition (VMD) (Dragomiretskiy and Zosso, 2013) is one of novel non-stationary signal decomposition techniques and has been recently applied in wind speed forecasting and many other fields (Dibaj et al., 2021; Yildiz et al., 2021). The original non-stationary signal can be decomposed into different band-limited intrinsic mode functions (IMFs) using VMD, similar to empirical mode decomposition (EMD). A non-recursive method is used to decompose original signal  $X(t)$  into  $M$  modes or subsequences  $u_m(m = 1, 2, \dots, M)$ . These IMFs have different central

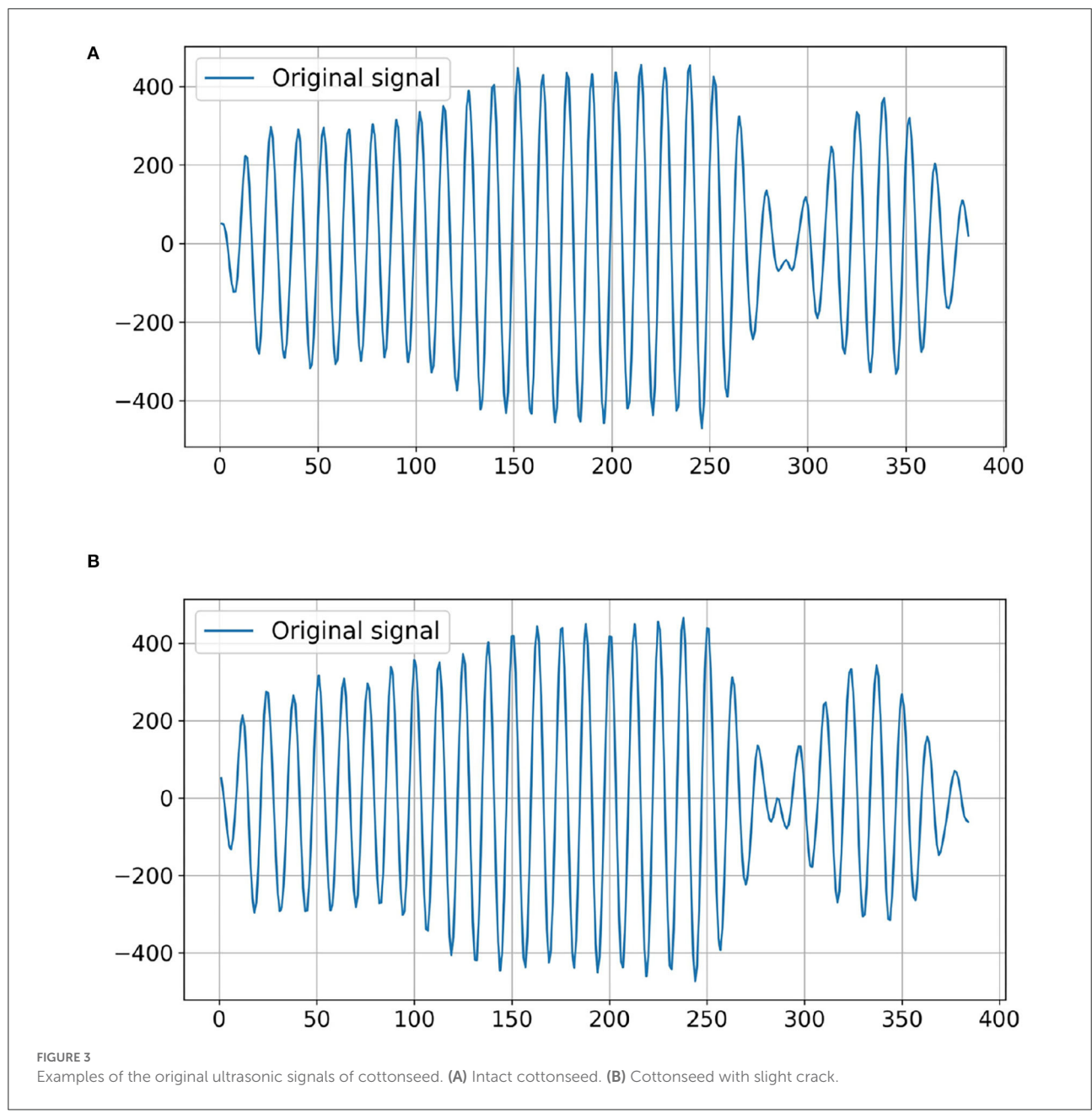
frequencies with finite bandwidths. The purpose of the transformation is to minimize the sum of the estimated frequency bandwidth of each IMF, and the constraint condition is that the sum of each IMF is equal to the input signal  $X(t)$ . The objective function and constraint condition corresponding to the variational constraint model can be represented by the following Equation 1:

$$\begin{cases} \min_{\{u_m\}, \{\omega_m\}} \sum_{m=1}^M \left\| \delta \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_m(t) e^{-j\omega_k t} \right] \right\|_2^2 \\ \text{s.t. } X(t) = \sum_{m=1}^M u_m \end{cases} \quad (1)$$

where  $\delta(t)$  is Dirac function;  $*$  denotes the convolution operation in signal processing;  $j$  is an imaginary number;  $\|\cdot\|$  denotes the  $L^2$ - norm;  $\{u_m\} = \{u_1, u_2, \dots, u_M\}$  is the set of all IMFs;  $\{\omega_m\} = \{\omega_1, \omega_2, \dots, \omega_M\}$  is the set of central frequencies corresponding to all IMFs.

In order to resolve the constrained variational optimization problem, quadratic penalty factor term  $\alpha$  and Lagrange multipliers  $\lambda(t)$  are defined, so the optimization problem can be transformed into an unconstrained variational problem. The constructed unconstrained form can be presented in Equation 2.

$$\begin{aligned} & \mathcal{L}(\{u_m\}, \{\omega_m\}, \{\lambda(t)\}) \\ &= \alpha \sum_{m=1}^M \left\| \delta \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_m(t) e^{-j\omega_k t} \right] \right\|_2^2 \\ &+ \left\| X(t) - \sum_{m=1}^M u_m \right\|_2^2 + \left\langle \{\lambda(t)\}, X(t) - \sum_{m=1}^M u_m \right\rangle \end{aligned} \quad (2)$$



where  $\alpha$  is used to ensure high reconstruction fidelity even in the presence of additive Gaussian white noise;  $\lambda(t)$  is used to strictly ensure the constraints.

To solve the unconstrained optimization problem in Equation 2, the alternating direction multiplier method (ADMM) (Hong and Luo, 2017) is used to find the saddle point of the optimization problem. The minimax point of the augmented Lagrangian function  $\mathcal{L}$  can be obtained by updating  $u_m$ ,  $\omega_m$  and  $\lambda$  alternately. The iteration process of  $u_m$ ,  $\omega_m$  and  $\lambda$  are as following

Equation 3, 4, 5:

$$\hat{u}_m^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq m} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_m)^2} \quad (3)$$

$$\omega_m^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_m(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_m(\omega)|^2 d\omega} \quad (4)$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau(\hat{X}(\omega) - \sum_m \hat{u}_m^{n+1}(\omega)) \quad (5)$$

where  $n$  is the number of iterations;  $\hat{X}(\omega)$ ,  $\hat{u}_m^{n+1}(\omega)$ ,  $\hat{u}_i(\omega)$  and  $\hat{\lambda}(\omega)$  can be obtained from  $X(t)$ ,  $u_m^{n+1}(t)$ ,  $u_i(t)$  and  $\lambda(t)$  by the Fourier transforms. If the convergence condition shown in Equation 6 is satisfied, the iteration will be terminated. Finally, through the inverse Fourier transform of  $\hat{u}_m^{n+1}(t)$ , the real part of the result is taken as the mode functions  $u_m^{n+1}(t)$ .

$$\sum_{m=1}^M \frac{\|\hat{u}_m^{n+1} - \hat{u}_m^n\|_2^2}{\|\hat{u}_m^n\|_2^2} < \epsilon \tag{6}$$

The optimization process for VMD is as follows:

- Step 1: Initialize  $u_m^1, \omega_m^1, \lambda^1$  and  $n$ , where  $n = 1$ .
- Step 2: Set  $n$  to  $n + 1$  and update  $\hat{u}_m^{n+1}(\omega)$ ,  $\omega_m^{n+1}$  and  $\hat{\lambda}^{n+1}$  according to Equation 3, 4, 5.
- Step 3: Repeat step 2 until the iteration convergence condition in Equation 6 is satisfied.  $M$  narrowband IMF  $u_m^{n+1}(t)$  can be obtained by using inverse Fourier transform.

The air-coupled ultrasonic signal of a cottonseed with slight crack is used as the examples for VMD decomposition. The results of VMD decomposition are shown in Figure 4. Because of the similarity between air-coupled ultrasonic signal of intact cottonseed and that of cottonseed with slight crack, it is important to extract the air-coupled ultrasonic signal features from these intrinsic mode functions.

### Encoding from ultrasound to image

The ultrasonic signal  $X(n)$  acquired from air-coupled ultrasonic detection system are decomposed to  $M$  intrinsic mode functions  $u_m, m \in \{1, 2, \dots, M\}$  using VMD transformation.  $M$  IMFs are stacked to generate the intrinsic mode functions matrix  $S$  of air-coupled ultrasonic signal  $X(n)$  according to Equation 7:

$$S = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix} = [s_1, s_2, \dots, s_L] \tag{7}$$

where  $S \in \mathbb{R}^{M \times L}$  and  $L$  is the length of the air-coupled ultrasonic signal  $X(n)$ .  $u_m$  is composed of  $L$  discrete values with the same length as the ultrasonic signal, which can be represented as  $u_m \in \mathbb{R}^{1 \times L}$ . Vector  $s_l$  is composed of  $M$  discrete values in the column direction of the matrix  $S$  which can be represented as  $s_l \in \mathbb{R}^{M \times 1}$ . In order to convert the ultrasonic signal  $X(n)$  into a color-coded image  $I_C$ , the color set  $C$  is defined for color coding,  $C = \{c_1, c_2, \dots, c_B\}$ , where  $c_i$  represents different colors and  $B$  is the number of different colors in the color set  $C$ . First, vector  $s_1$  at the first column of the intrinsic mode functions matrix  $S$  is selected and converted into a part

of colorful image  $I_C$ . Here,  $s_1 = \{u_{11}, u_{12}, \dots, u_{1M}\}$ . Color  $c_1$  in set  $C$  is chosen and used to draw a polyline in the image  $I_C$  according to the value of  $s_1$ . Then vector  $s_2$  at the second column of matrix  $S$  and color  $c_2$  are selected to draw the second polyline in the image  $I_C$  to represent  $s_2$ . According to this method, color  $c_1$  is used to draw the next polyline again after completing the drawing of  $B$  colorful polylines. Finally,  $B$  is considered as the cycle to complete the drawing of  $L$  colorful polylines. Image  $I_C$  is generated from the air-coupled ultrasonic signal  $X(n)$  using the above color encoding method.

In order to accurately convert the ultrasonic signal into a colorful image, first, the drawing ranges  $W_S, H_S$  in image  $I_C$  and the coordinates  $(x_{O_S}, y_{O_S})$  of the image origin  $O_S$  are defined. Then the coordinates required to draw polylines are calculated according to Equation 8:

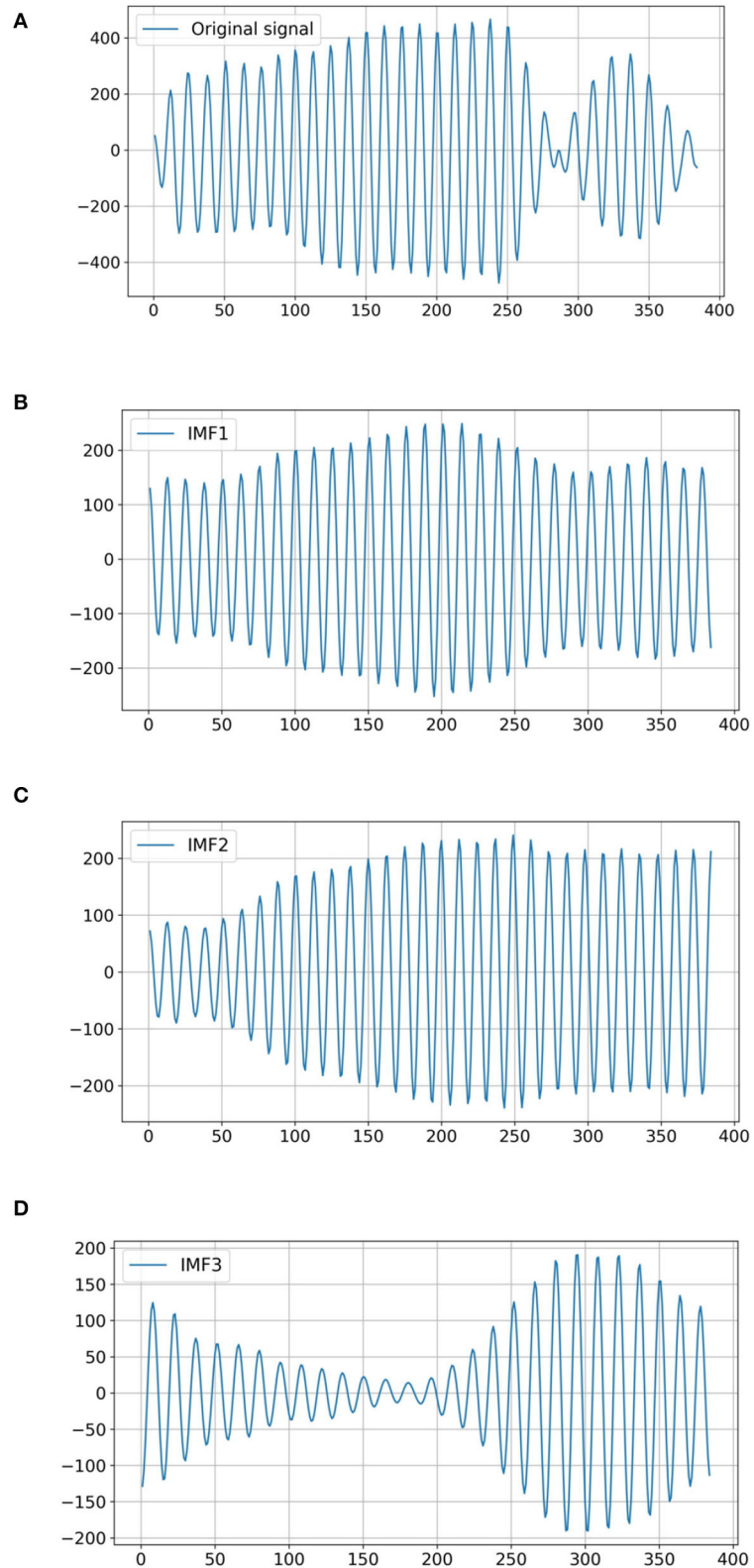
$$\begin{cases} x_{s_{lm}} = x_{O_S} + \frac{(m-1)W_S}{M-1} & m \in \{1, 2, \dots, M\} \\ y_{s_{lm}} = y_{O_S} + \frac{(H_S - y_{O_S})(S_{\max} - S_{lm})}{(S_{\max} - S_{\min})} & l \in \{1, 2, \dots, L\} \end{cases} \tag{8}$$

where  $S_{\max}$  and  $S_{\min}$  are respectively the maximum and minimum values in the matrix  $S$  composed of  $M$  intrinsic mode functions after VMD decomposition of each ultrasonic signal. The position of each column vector of matrix  $S$  in the image  $I_C$  is determined by Equation 8, and then the corresponding position is connected with the same color to complete the drawing of polylines. For each vector  $s_l$ ,  $M - 1$  colorful lines need to be drawn. In this study, let  $M = 3, L = 450, B = 10$ , the size of the generated image  $I_C$  is  $1,800 \times 1,200$  and the origin coordinate of the encoding image part on image  $I_C$  is  $(288, 1010)$ . The process of generating images from the ultrasonic signals of intact cottonseed and slight cracked cottonseed is shown in Figure 5. As seen in Figure 5, the coding representation of ultrasonic signals from one-dimensional ultrasonic signal to two-dimensional images can be realized by polylines with different colors according to the results of VMD decomposition of ultrasonic signals.

In practical application, in order to obtain the optimal detection effect, the air-coupled ultrasonic data of 30 intact cottonseeds and 30 cottonseeds with slight crack are randomly selected from the air-coupled ultrasonic data set, the background image is generated according to the above method, and then the air-coupled ultrasonic data of the remaining 540 cottonseeds are used to generate colorful encoding images on the background image. So, the conversion from sound to image is completed. Finally, the image data set including 266 images generated from intact kernels and 274 images generated from slight crack kernels is obtained. In this study, 80% of the data set is used to train the model and 20% is used to test the effect of the training model.

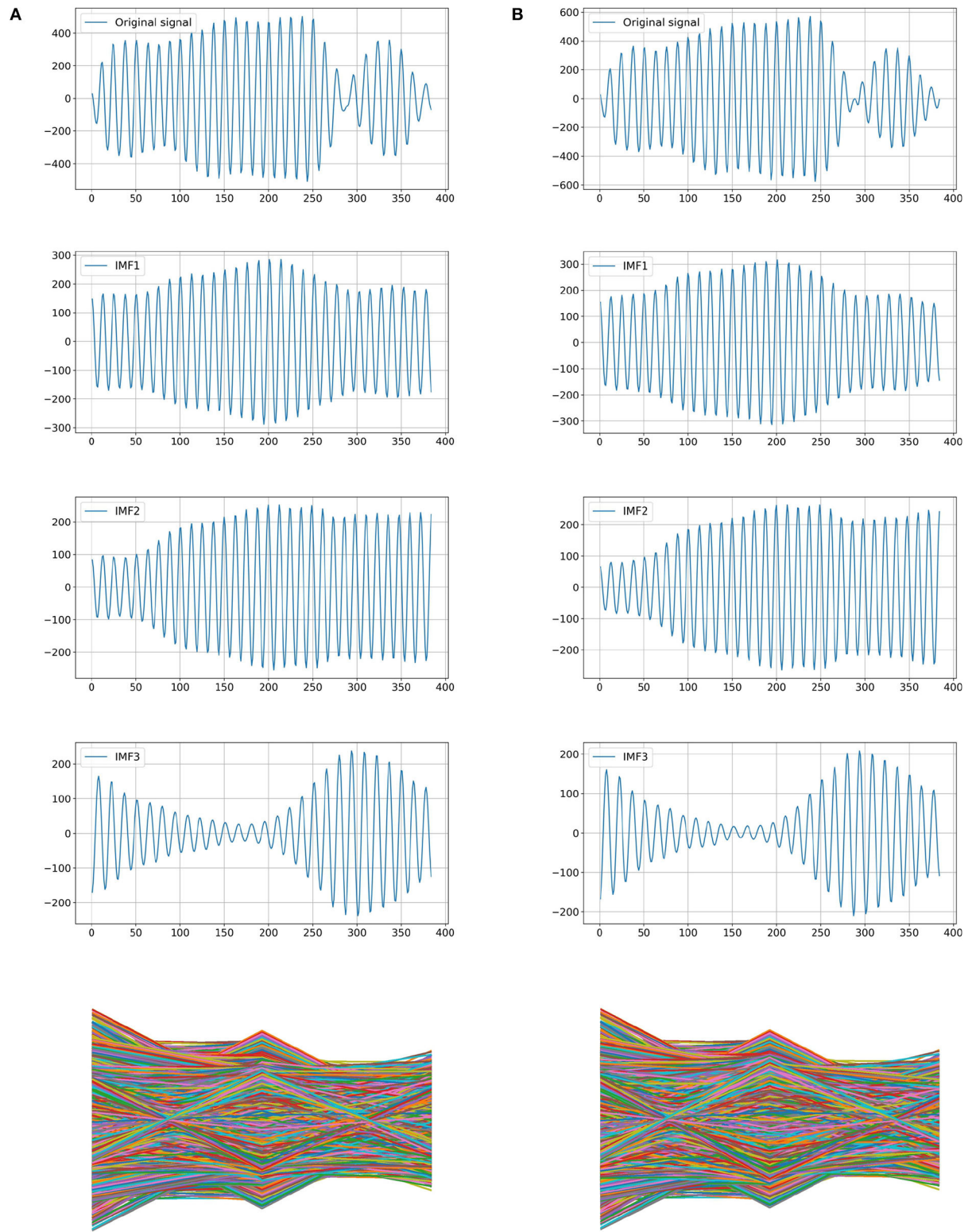
### MobileViT vision transformer model

The transformer-based model achieves great success in the natural language processing (NLP) field (Vaswani et al., 2017).



**FIGURE 4**  
The VMD decomposition from the air-coupled ultrasonic signal of a cottonseed with slight crack. **(A)** Original signal. **(B)** IMF1. **(C)** IMF2. **(D)** IMF3.





**FIGURE 5**  
Colorful images generated from the ultrasonic signals. **(A)** The process of generating image from the ultrasonic signal of intact cottonseed. **(B)** The process of generating image from the ultrasonic signal of slightly cracked cottonseed.

Transformer has a layer stacking architecture, which only uses the multi-Head self-attention mechanism without convolution and recursion. Inspired by the successful application of transformer in NLP field, Dosovitskiy et al. (2020) propose Vision Transformer (ViT) model which employed a standard Transformer directly to visual tasks. In this method, the image is divided into a sequence of patches, and the linear embedding sequence of these image patches are taken as the input of the Transformer. The processing method of image patches is the same as that of tokens in NLP application, and excellent results are achieved on massive data sets. Liu et al. (2021), proposed the Swin Transformer (Shifted Window Transformer) model that could replace the classic convolutional neural network (CNN) architecture and become a general backbone in the field of computer vision. This model is based on the idea of ViT model and shows the effectiveness on different vision problems using patch merging and the shifted window with self-attention.

Although Vision Transformer-based model can be an alternative to CNNs in computer vision field, the large model size, high requirement for training data, and latency of Vision Transformer limits its practice application, especially for resource constrained equipment. In order to obtain a lightweight and efficient architecture of Vision Transformer model, MobileViT (Mehta and Rastegari, 2021) is proposed for mobile vision applications. MobileViT combines the advantages of transformers and convolutions, so it can encode the local information obtained from convolutions and global information obtained from transformers in tensors without lacking in inductive bias. The model structure of MobileViT is shown in Figure 6, where MV2 block represents MobileNetV2 block with inverted residual structure.  $\downarrow 2$  refers to down-sampling operation. A standard  $3 \times 3$  convolution operation is represented by  $Conv-3 \times 3$  block in MobileViT model. MobileViT block combines convolutions with transformers to learn the global and local information from input tensor respectively. Tensor  $X_T \in \mathbb{R}^{H \times W \times D}$  passes through a series of convolution operations and tensor  $X_{L_T} \in \mathbb{R}^{H \times W \times d}$  is obtained. Then tensor  $X_{L_T}$  is divided into  $N$  patches with width  $w$  and height  $h$  and  $X_{U_T} \in \mathbb{R}^{P \times N \times d}$  can be obtained by unfolding  $X_{L_T}$ . For example, the  $X_{L_T}$  in Figure 6 is equally divided into  $4 \times 4$  small patches, then a  $d$ -dimensional vector is extracted at the same position of each patch. The corresponding vectors from the same position  $p \in \{1, \dots, P\}$  are combined and used to generate tokens. The inter-path relationships  $X_{G_T} \in \mathbb{R}^{P \times N \times d}$  can be obtained by encoding operation with Transformers as Equation 9:

$$X_{G_T} = \text{Transformer}(X_{U_T}(p)), 1 \leq p \leq P \quad (9)$$

where local information can be encoded by  $X_{U_T}(p)$  and global information passing through

$p$ -th position in  $P$  patches can be encoded by  $X_{U_T}(p)$ . The output  $Y_T$  of MobileViT block can be obtained after convolution and concatenation operations.

The computation of self-attention in above Transformer is realized using scale dot-product attention according to Equation 10:

$$\text{Attention}(Q, K, V) = \text{SoftMax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (10)$$

where  $Q$ ,  $K$  and  $V$  represent the query, key, and value matrices respectively.  $d_k$  refers to the dimension of  $Q$  and  $K$ . The *SoftMax* function is used to obtain the weights of  $V$ . Spatial order of pixels will be ignored in standard ViTs. But both the spatial order of pixels and the path order will be obtained in MobileViT model.

## Results and discussion

### Cottonseed germination test

It is important to guarantee the safety of air-coupled ultrasonic detection for cottonseed. In order to verify the safety of ultrasonic detection, the germination test of cottonseeds is used as the verification method. First, two batches of intact cottonseeds are randomly selected, then the cottonseeds are divided into two groups. One group is not used to ultrasonic testing, and the other group is passed through air-coupled ultrasonic at the same frequency and intensity in detection experiment. For each germination test, 1,000 cottonseeds after ultrasonic testing and 1,000 cottonseeds without ultrasonic testing are selected respectively. Then the two groups of cottonseeds are used in cottonseed germination tests at the same time. It is determined whether air-coupled ultrasonic detection will cause damage to the cottonseeds by the comparison of the germination rates.

Table 1 shows the comparative results of the cottonseed germination rate. By comparing the results in Table 1, it can be seen that air-coupled ultrasonic detection does not affect the germination rate of cottonseeds. At the same time, the germination rate of seeds tends to decrease with the increase of time. This trend may be due to the storage environment of cottonseeds not in accordance with the storage standards, which affects the germination rate of cottonseeds. Therefore, in this study, the air-coupled ultrasound with frequency of 400 kHz and voltage of 200 V is used to detect the slight crack of cottonseed, which will not damage the cottonseeds or reduce the germination rate of cottonseeds. This is a safe non-destructive detection method.

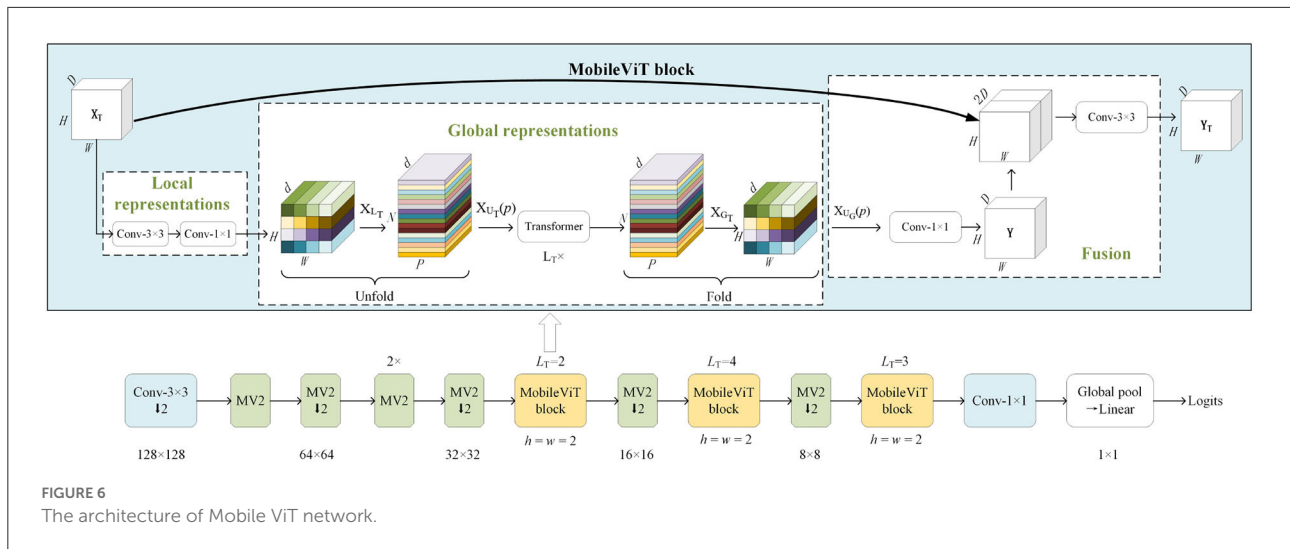


FIGURE 6 The architecture of Mobile ViT network.

TABLE 1 The results of cottonseeds germination test.

Cottonseed category	Storage for 2 months	Storage 1 year
Cottonseeds after ultrasonic testing	83.5%	72.2%
Cottonseeds without ultrasonic testing	86.1%	70.6%

### Determination of decomposition number *M* using sample entropy

The ultrasonic signal  $X(n)$  is decomposed to  $M$  intrinsic mode functions using VMD decomposition. The number  $M$  of intrinsic mode functions needs to be determined artificially. Sample entropy is proposed to measure the complexity of time series by Richman and Moorman (2000). Complex time series signals have large sample entropy, while time series signals with strong selfsimilarity have small sample entropy. Zhang et al. (2020) proposed a method to select the number  $M$  of intrinsic mode functions by computing the sample entropy of time series. For an appropriate  $M$  value of VMD decomposition operation, it will correspond to smaller sample entropy of the time sequence signal. Therefore, on the condition of effectively limiting the computational complexity, the  $M$  value is determined by calculating the sample entropy of the intrinsic mode functions.

For one-dimensional time series signal  $\{X_{SE}(i), i = 1, 2, \dots, L\}$ ,  $z$ -dimensional vector is reconstructed and represented by  $\{Y_{SE}(i), i = 1, 2, \dots, Z, Z = L - z + 1\}$  according to Equation 11:

$$Y_{SE}(i) = \{X_{SE}(i), X_{SE}(i + 1), X_{SE}(i + 2), \dots, X_{SE}(i + m - 1)\} \quad (11)$$

Then the maximum value of Euclidean distance between any component of vectors  $Y_{SE}(i)$  and  $Y_{SE}(j)$  is calculated according

to Equation 12 and represented as  $D_{SE}(Y_{SE}(i), Y_{SE}(j))$ .

$$D_{SE}(Y_{SE}(i), Y_{SE}(j)) = \max [|Y_{SE}(i + k) - Y_{SE}(j + k)|] \quad (12)$$

where  $i, j \in \{1, 2, \dots, Z - z + 1\}$  and  $k \in \{0, 1, \dots, z - 1\}$ . Then  $A_i^z(r)$  is calculated according to Equation 13, where  $r$  is tolerance threshold and  $A_i^z$  is the number that the distance between  $Y_{SE}(i)$  and  $Y_{SE}(i)$  is not greater than  $r$ .

$$\bar{A}_i^z(r) = \frac{1}{L - z + 1} A_i^z \quad (13)$$

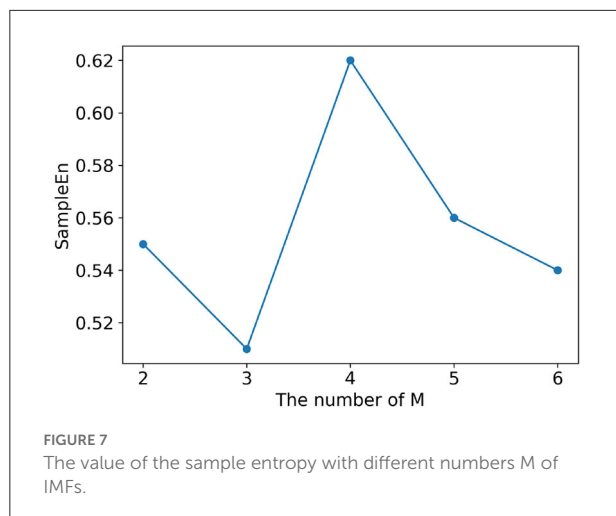
Then  $A_i^z(r)$  is used to calculate  $\bar{A}_i^z(r) \cdot \bar{A}_i^z(r) = \frac{1}{L - z} \sum_{i=1}^{L-z} A_i^z(r)$ . Finally, the sample entropy of time series signal is calculated according to the following to Equation 14:

$$SampleEn(m, r) = -\ln \frac{A_i^{z+1}(r)}{A_i^z(r)} \quad (14)$$

The original air-coupled ultrasonic signal is decomposed using VMD with different parameter  $M$ . Then the sample entropy corresponding to each  $M$  is calculated accordance to Equation 14. In order to obtain an appropriate detection speed,  $M$  is set to 2, 3, 4, 5, 6 in the experiment respectively. It can be seen from Figure 7 that the sample entropy has the smallest value when  $M$  is equal to 3. Therefore,  $M = 3$  is determined as the number of intrinsic mode functions of VMD decomposition in this study.

### Influence of the number of encoding colors

The conversion from air-coupled ultrasonic data to image data is realized through a specific number of color encoding sets. In order to determine the optimal number of encoding



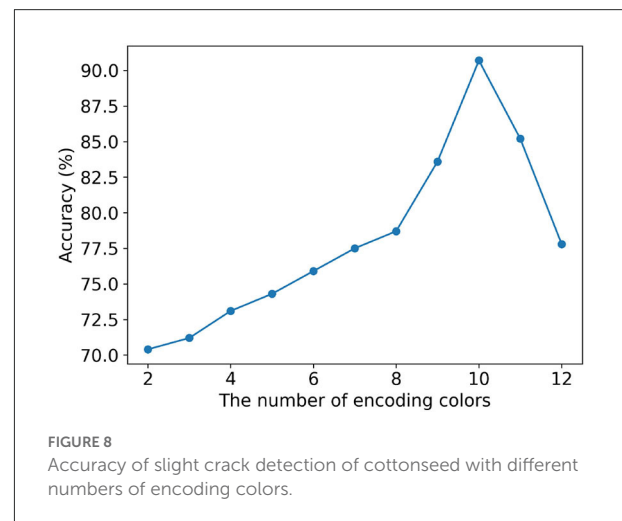
colors for ultrasound to image conversion, a color set for encoding including 12 colors is constructed. The air-coupled ultrasonic signal decomposed by VMD method is encoded by two to 12 colors respectively. Eleven image data sets are constructed using the original air-coupled ultrasonic data. The MobileViT model is trained by using each image data set. The training data sets and test data sets from 11 image data sets are divided in the same proportion respectively. The initialization parameters of MobileViT in this study are shown in [Table 2](#). These initialization parameters are mainly determined according to GPU performance and characteristics of encoding image. The comparison results of the MobileViT model obtained on the test sets of colorful images generated by different numbers of encoding colors are shown in [Figure 8](#). The colorful image sets for tests come from the same air-coupled ultrasonic data test set, but different test sets are generated according to the corresponding color encoding method. It can be seen from [Figure 8](#) that with the increase of encoding colors, the classification accuracy increases gradually. When the number of encoding colors is equal to 10, the classification accuracy reaches the maximum. Then the classification accuracy decreases with the increase of the encoding color number. This is mainly because, with the increase of the encoding colors number and the types of colors, the image generated from the air-coupled ultrasound becomes more complex. It is difficult for the classification algorithm to extract effective features from the complex image. Therefore, to realize the conversion from sound to image, 10 colors are selected to encode the intrinsic mode functions of the original air-coupled ultrasonic signal.

## Comparison of different methods

In order to compare the detection effects of the method proposed in this study with those of other methods, eight

**TABLE 2** Parameters of MobileViT for cottonseed with slight crack.

Parameters	Value
Input size	256 × 256
Classes	2
Batch size	16
Learning rate	$1.0 \times 10^{-3}$
Iterations	10



detection methods are proposed for the comparison experiment. The long short-term memory (LSTM) network is a special kind of recurrent neural network that can store and retrieve information from sequence data by memory cells ([Hochreiter and Schmidhuber, 1997](#)). LSTM network is an effective classifier for dealing with one-dimensional time-series and sequential data. The original air-coupled ultrasonic signal, wavelet features, and the results of VMD decomposition are used as input data respectively, and then combined with LSTM classifier to obtain four methods for the comparison. The wavelet transformer can be used to obtain the time-frequency domain features of original air-coupled ultrasonic signal.

The color images used to train deep learning classifiers are shown in [Figure 9](#). For the data set obtained by the color encoding (CE) method from sound to image, different classifiers are used to compare the detection of cottonseed with slight crack. The traditional convolutional neural network (CNN) ([Fan S. et al., 2020](#)), residual network (ResNet18) ([He et al., 2016](#)) and Swin Transformer ([Liu et al., 2021](#)) are chosen as the classifier for the comparison respectively. The shortcut connections in ResNet18 make deeper neural networks to realize complex classification tasks. These models use a cross-entropy function as a loss function. Meanwhile, the training iterations of all models is set to 10.

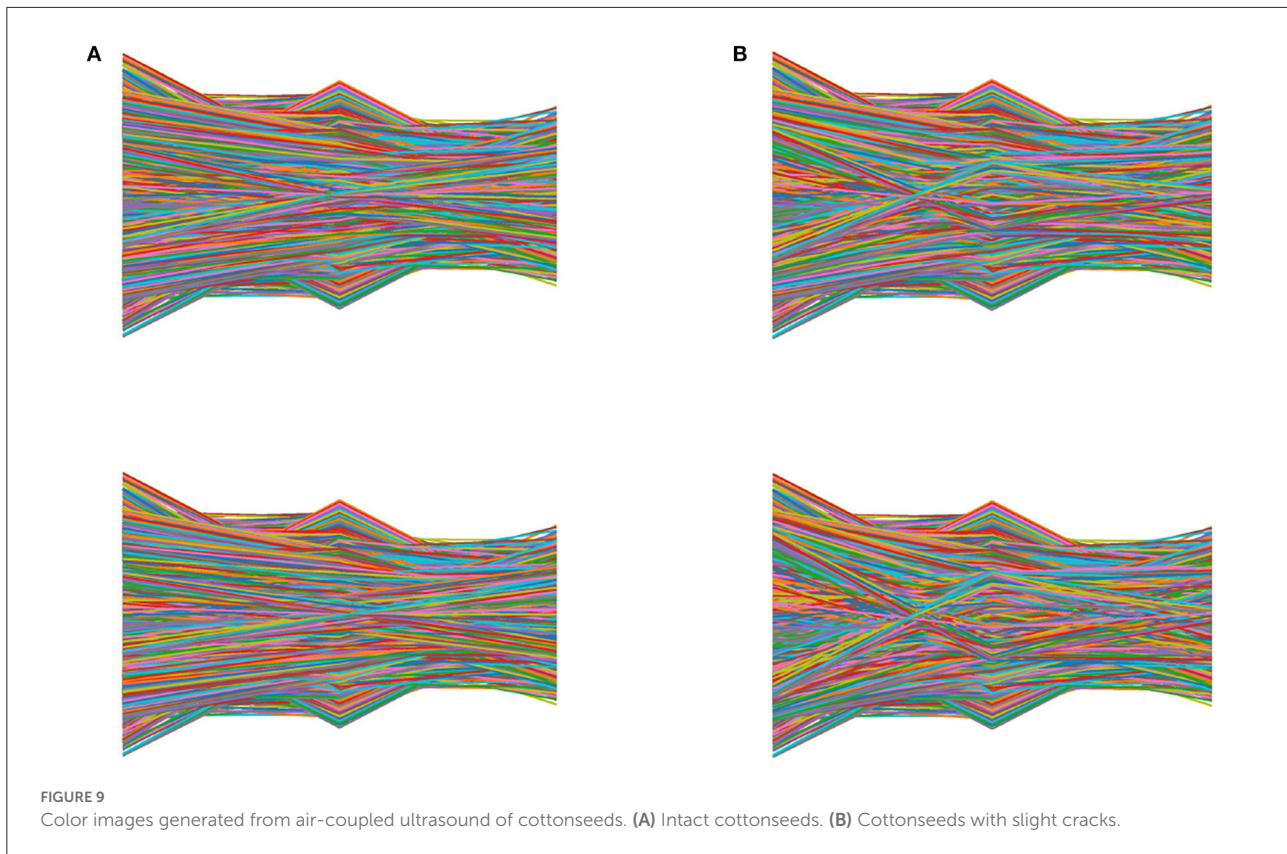


TABLE 3 The comparison of different methods for the detection of cottonseed with slight crack.

Method	Precision	Recall	F1 score	Accuracy
1D raw data - LSTM	80%	65.5%	72%	74%
Wavelet - LSTM	64.1%	92.6%	78.5%	70.4%
VMD- LSTM	75.5%	68.5%	71.8%	73.1%
VMD- CE - CNN	94.4%	61.8%	74.7%	78.7%
VMD- CE -	82.8%	96.4%	89.1%	88%
ResNet18				
VMD- CE - Swin	73.1%	89.1%	80.3%	77.8%
Transformer				
VMD- CE -	86.9%	96.4%	91.4%	90.7%
MobileViT (proposed method)				

The experimental results are shown in Table 3. From the experimental results, it can be seen that the two-dimensional image can better reflect the features of air-coupled ultrasonic signal than the features of one-dimensional signal. Because the MobileViT model can combine the local analysis ability of convolutional neural network with the global analysis ability of Transformer, the classification performance of MobileViT

is better than that of convolutional neural network and Swin Transformer model. Due to the vast amount of parameters in the Swin Transformer model and the requirement for a large number of training data, the training and learning of non-large data sets can't show the advantages of the Swin Transformer model. Therefore, VMD-CE-MobileViT approach proposed in this study can distinguish normal cottonseed from cottonseed with slight crack effectively.

## Conclusion

In this paper, a detection method based on air-coupled ultrasound and sound to image encoding is proposed for slight crack identification of cottonseed. The traditional ultrasound detection method is not suitable for the requirements of non-destructive detection of cottonseed quality, and it is very difficult to detect cottonseed with slight crack using machine vision and other non-destructive detection technologies. To distinguish kernels with slight crack from intact kernels, a non-destructive, non-contact detection method based on air-coupled ultrasound is developed. VMD decomposition is used to obtain the IMFs of air-coupled ultrasonic signal. Then the feature matrix from the IMFs is applied to generate colorful image by a color encoding method. This method of converting sound into image can help

the MobileViT classifier to obtain higher detection accuracy. The experimental results show that slightly cracked cottonseed can be distinguished from normal cottonseed precisely. The average accuracy of slightly cracked cottonseed identification test is 90.7%. The presented method can be extended to other signal recognition domain, such as distinguishing premature heartbeat signal from normal heartbeat signal in electrocardiogram (ECG) domain. In the future, we will combine the method proposed in this study with hyperspectral image processing technology to improve the detection accuracy further.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

## Author contributions

CZ proposed the idea, designed the methodology and prepared the manuscript. WH managed and coordinated the research activity planning and execution. XL conducted the experiment and maintained the research data. XH designed the experimental system. XT assisted in analyzing experimental data. LC provided oversight and leadership responsibility for the research. QW reviewed, edited the manuscript, and is responsible for ensuring that the descriptions are accurate and

agreed by all authors. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Natural Science Foundation of China (NSFC No. 31871523), Young Elite Scientists Sponsorship Program by CAST (2019QNRC001), and National Natural Science Foundation of China (NSFC No. 31901402).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Baek, I., Kim, M. S., Cho, B.-K., Mo, C., Barnaby, J. Y., et al. (2019). Selection of optimal hyperspectral wavebands for detection of discolored, diseased rice seeds. *Applied Sciences* 9, 1027. doi: 10.3390/app9051027
- Bai, X., Wang, W., Jiang, H., Chu, X., Zhao, X., Wang, B., et al. (2018). "Research on classification method of cotton seeds on machine vision," in: *2018 ASABE Annual International Meeting* (American Society of 385 Agricultural and Biological Engineers), Detroit, 1. doi: 10.13031/aim.201800810
- Barboza da Silva, C., Oliveira, N. M., de Carvalho, M. E. A., de Medeiros, A. D., de Lima Nogueira, M., and Dos Reis, A. R. (2021). Autofluorescence-spectral imaging as an innovative method for rapid, non-destructive and reliable assessing of soybean seed quality. *Sci. Rep.* 11, 17834. doi: 10.1038/s41598-021-97223-5
- Belin, E., Douarre, C., Gillard, N., Franconi, F., Rojas-Varela, J., Chapeau-Blondeau, F., et al. (2018). Evaluation of 3d/2d imaging and image processing techniques for the monitoring of seed imbibition. *J. Imaging* 4, 83. doi: 10.3390/jimaging4070083
- Chelladurai, V., Karuppiah, K., Jayas, D., Fields, P., and White, N. (2014). Detection of *Callosobruchus maculatus* (f.) infestation in soybean using soft X-ray and NIR hyperspectral imaging techniques. *J. Stored Prod. Res.* 57, 43–48. doi: 10.1016/j.jspr.2013.12.005
- Dibaj, Ali, Etefagh, M. M., Hossain, R., and Eghaghi, M. B. (2021). A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults. *Expert Syst. Appl.* 167, 114094. doi: 10.1016/j.eswa.2020.114094
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). An image is worth 16x16 words: transformers for image recognition at scale. *arXiv*. preprint arXiv:2010.11929. doi: 10.48550/arXiv.2010.11929
- Dragomiretskiy, K., and Zosso, D. (2013). Variational mode decomposition. *IEEE Trans. Signal Process.* 62, 531–544. doi: 10.1109/TSP.2013.2288675
- ElMasry, G., ElGamal, R., Mandour, N., Gou, P., Al-Rejaie, S., Belin, E., et al. (2020). Emerging thermal imaging techniques for seed quality evaluation: principles and applications. *Food Res. Int.* 131, 109025. doi: 10.1016/j.foodres.2020.109025
- Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., et al. (2020). On line detection of defective apples using computer vision system combined with deep learning methods. *J. Food Eng.* 286, 110102. doi: 10.1016/j.jfoodeng.2020.110102
- Fan, Y., Ma, S., and Wu, T. (2020). Individual wheat kernels vigor assessment based on NIR spectroscopy coupled with machine learning methodologies. *Infrared Phys. Technol.* 105, 103213. doi: 10.1016/j.infrared.2020.103213
- Fang, W., Lu, W., Xu, H., Hong, D., and Liang, K. (2016). Study on the detection of rice seed germination rate based on infrared thermal imaging technology combined with generalized regression neural network. *Spectrosc. Spectr. Anal.* 36, 2692–2697. doi: 10.3964/j.issn.1000-0593(2016)08-2692-06
- Fang, Y., Lin, L., Feng, H., Lu, Z., and Emms, G. W. (2017). Review of the use of air-coupled ultrasonic technologies for nondestructive testing of wood and wood products. *Comput. Electr. Agric.* 137, 79–87. doi: 10.1016/j.compag.2017.03.015
- Fariñas, L., Contreras, M., Sanchez-Jimenez, V., Benedito, J., and Garcia-Perez, J. V. (2021a). Use of air-coupled ultrasound for the non-invasive characterization of the textural properties of pork burger patties. *J. Food Eng.* 297, 110481. doi: 10.1016/j.jfoodeng.2021.110481
- Fariñas, L., Sanchez-Torres, E. A., Sanchez-Jimenez, V., Diaz, R., Benedito, J., Garcia-Perez, J. V., et al. (2021b). Assessment of avocado textural changes during

- ripening by using contactless air-coupled ultrasound. *J. Food Eng.* 289, 110266. doi: 10.1016/j.foodeng.2020.110266
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 770–778. doi: 10.1109/CVPR.2016.90
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Hong, M., and Luo, Z.-Q. (2017). On the linear convergence of the alternating direction method of multipliers. *Math. Program.* 162, 165–199. doi: 10.1007/s10107-016-1034-2
- Hu, N., Li, W., Du, C., Zhang, Z., Gao, Y., Sun, Z., et al. (2021). Predicting micronutrients of wheat using hyperspectral imaging. *Food Chem.* 343, 128473. doi: 10.1016/j.foodchem.2020.128473
- Huang, M., Wang, Q., Zhu, Q., Qin, J., and Huang, G. (2015). Review of seed quality and safety tests using optical sensing technologies. *Seed Sci. Technol.* 43, 337–366. doi: 10.15258/sst.2015.43.3.16
- Huang, S., Fan, X., Sun, L., Shen, Y., and Suo, X. (2019). Research on classification method of maize seed defect based on machine vision. *J. Sens.* 2019, 1–9. doi: 10.1155/2019/2716975
- Kim, G., Kim, G., H., Lohumi, S., Kang, J., S., Cho, B., et al. (2014). Viability estimation of pepper seeds using time-resolved photothermal signal characterization. *Infrared Phys. Technol.* 67, 214–221. doi: 10.1016/j.infrared.2014.07.025
- Li, D., and Liu, Y., and Gao, L. (2016). Research of maize seeds classification recognition based on the image processing. *Int. J. Signal Proces. Image Proces. Pattern Recognit.* 9, 181–190. doi: 10.14257/ijsp.2016.9.11.16
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). “Swin transformer: Hierarchical vision transformer using shifted windows,” in: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, 10012–10022. doi: 10.1109/ICCV48922.2021.00986
- Mehta, S., and Rastegari, M. (2021). Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer. *arXiv*. preprint arXiv:2110.02178. doi: 10.48550/arXiv.2110.02178
- Pearson, T. (2001). Detection of pistachio nuts with closed shells using impact acoustics. *Appl. Eng. Agric.* 17, 249. doi: 10.13031/2013.5450
- Pearson, T. C., Cetin, A. E., Tewfik, A. H., and Haff, R. P. (2007). Feasibility of impact-acoustic emissions for detection of damaged wheat kernels. *Digit. Signal Process.* 17, 617–633. doi: 10.1016/j.dsp.2005.08.002
- Rahman, A., and Cho, B.-K. (2016). Assessment of seed quality using non-destructive measurement techniques: a review. *Seed Sci. Res.* 26, 285–305. doi: 10.1017/S0960258516000234
- Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–49. doi: 10.1152/ajpheart.2000.278.6.H2039
- Rodríguez-Pulido, F. J., and Ferrer-Gallego, R., González-Miret, M. L., Rivas-Gonzalo, J. C., Escribano-Bailón, M. T., and Heredia, F. J. (2012). Preliminary study to determine the phenolic maturity stage of grape seeds by computer vision. *Anal. Chim. Acta* 732, 78–82. doi: 10.1016/j.aca.2012.01.005
- Sun, X., Guo, M., Ma, M., and Mankin, R. W. (2018). Identification and classification of damaged corn kernels with impact acoustics multi-domain patterns. *Comput. Electron. Agric.* 150, 152–161. doi: 10.1016/j.compag.2018.04.008
- Sunoj, S., Igathinathane, C., and Visvanathan, R. (2016). Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. *Comput. Electron. Agric.* 124, 234–242. doi: 10.1016/j.compag.2016.04.012
- Tao, F., Yao, H., Hruska, Z., Liu, Y., Rajasekaran, K., Bhatnagar, D., et al. (2019). Use of visible-near-infrared (Vis-NIR) spectroscopy to detect aflatoxin B1 on peanut kernels. *Appl. Spectrosc.* 73, 415–423. doi: 10.1177/0003702819829725
- Tiitta, M., Tiitta, V., Gaal, M., Heikkinen, J., Lappalainen, R., Tomppo, L., et al. (2020). Air-coupled ultrasound detection of natural defects in wood using ferroelectric and piezoelectric sensors. *Wood Sci. Technol.* 54, 1051–1064. doi: 10.1007/s00226-020-01189-y
- Tu, K.-I., Li, L.-j., Yang, L.-m., Wang, J.-h., and Qun, S. (2018). Selection for high quality pepper seeds by machine vision and classifiers. *J. Integr. Agric.* 17, 1999–2006. doi: 10.1016/S2095-3119(18)62031-3
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). “Attention is all you need,” in: *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA 30, 5998–6008.
- Xia, Y., Xu, Y., Li, J., Zhang, C., and Fan, S. (2019). Recent advances in emerging techniques for non-destructive detection of seed viability: A review. *Artif. Intell. Agri.* 1, 35–47. doi: 10.1016/j.iaia.2019.05.001
- Yang, G., Wang, Q., Liu, C., Wang, X., Fan, S., Huang, W., et al. (2018). Rapid and visual detection of the main chemical compositions in maize seeds based on Raman hyperspectral imaging. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 200, 186–194. doi: 10.1016/j.saa.2018.04.026
- Yanyun, J., Wanlin, G., Han, Z., Dong, A., Sihan, G., Ahmed, S. I., et al. (2016). Identification of damaged corn seeds using air-coupled ultrasound. *Int. J. Agric. Biol. Eng.* 9, 63–70. doi: 10.3965/ij.ijabe.20160901.1880
- Yildiz, C., Acikgoz, H., Korkmaz, D., and Budak, U. (2021). An improved residual-based convolutional neural network for very short-term wind power forecasting. *Energy Convers. Manage.* 228, 113731. doi: 10.1016/j.enconman.2020.113731
- Zhang, Y., and Guo, W. (2020). Moisture content detection of maize seed based on visible/near-infrared and near-infrared hyperspectral imaging technology. *Int. J. Food Sci. Technol.* 55, 631–640. doi: 10.1111/ijfs.14317
- Zhang, Y., Pan, G., Chen, B., Han, J., Zhao, Y., Zhang, C., et al. (2020). Short-term wind speed prediction model based on GA-ANN improved by VMD. *Renew. Energy* 156, 1373–1388. doi: 10.1016/j.renene.2019.12.047
- Zhou, Q., Huang, W., Fan, S., Zhao, F., Liang, D., Tian, X., et al. (2020). Non-destructive discrimination of the variety of sweet maize seeds based on hyperspectral image coupled with wavelength selection algorithm. *Infrared Phys. Technol.* 109, 103418. doi: 10.1016/j.infrared.2020.103418