RESEARCH ARTICLE

Deep learning-based algorithm for automatic identification and classification of surface damage of agricultural products

Weili Liu*

School of Electronic Information Engineering, City University of Zhengzhou, Zhengzhou, Henan, China.

Received: December 19, 2023; **accepted:** April 1, 2024.

Traditional surface damage detection algorithms for agricultural products cannot be applied to large-scale instances due to the difficulty of their implementation. Thus, research on automatic identification and classification algorithms for agricultural products surface damage based on deep learning has emerged. This study proposed a deep learning-based algorithm for automatic identification and classification of surface damage of agricultural products and compared it with six mainstream classification models in plantvillage in terms of generalization ability, training time, and amount of pre-training data. The results proved that the model proposed in this study was the best in all aspects and had the highest accuracy when the pre-training data reached 40,000. This study verified the superiority and generalization ability of the proposed model and provided a new solution and reference standard in the field of agricultural product surface damage detection. Further, the results provided valuable reference and inspiration for related research and practice.

Keywords: deep learning; agricultural products; surface damage; automatic identification; classification.

***Corresponding author:** Weili Liu, School of Electronic Information Engineering, City University of Zhengzhou, Zhengzhou 452370, Henan, China. Emails: liu_weili@outlook.com.

Introduction

The maintenance of the quality of agricultural products is crucial, especially in the accurate identification and classification of their surface damage, which is significant for ensuring food safety and reducing economic losses. Currently, although the surface damage detection of agricultural products is widely used by manual detection means, this method cannot meet the requirements of large-scale, rapid, and accurate detection in the modern agricultural supply chain due to its strong subjectivity, low efficiency, limited accuracy, and high cost [1].

The application of deep learning techniques in the field of agricultural product surface damage recognition has become an emerging trend. Deep learning architectures such as convolutional neural networks (CNNs) [2], long and short-term memory networks (LSTMs) [3], and multilayer perceptual machines (MLPs) [4] have been successfully applied to image feature extraction and complex classification tasks, which significantly improve the recognition accuracy and system robustness. Among them, CNNs have been widely used in image feature extraction [5], while LSTMs and MLPs have demonstrated strong performance in complex sequence data and classification tasks [6, 7]. However, existing techniques still face some challenges, such as the difficulty in effectively extracting and distinguishing injury site features, and the limited ability to adapt to different kinds of agricultural products and changes in injury types [8-10].

The core objective of this study was to develop an efficient, accurate, and generalized automatic identification and classification system based on computer vision and deep learning techniques to address the many challenges of agricultural product surface damage recognition. The system aimed to overcome the inherent defects of traditional manual detection methods and realize the refined identification and classification of surface damage of agricultural products in different kinds, states, and viewpoints by deeply integrating deep learning techniques such as CNN, attention mechanism, LSTM, and MLP to substantially improve the efficiency and accuracy of agricultural product quality monitoring, reduce economic losses due to surface damage, protect food safety, and positively promote the development of intelligent detection technology in the agricultural industry. By constructing and validating a large image database containing multiple agricultural product samples and damage types, this study would verify the effectiveness and superiority of the proposed algorithm in real-world complex scenarios. This study designed a feature extraction method based on CNN and attention mechanism to learn and highlight damage features from a huge agricultural product image dataset in an automated way, weakening the background noise and non-relevant feature interference. Then, a multi-label classification framework was developed, which combined LSTM and MLP in order to realize the simultaneous and accurate determination of damage category, degree, and location by integrating the information of produce type, state, and angle. The proposed algorithm was verified by using a diverse and large-scale image database including a wide range of common produce types (e.g., apples, tomatoes, corn, potatoes, *etc.*) and damage types (cracks, spots, pests, and diseases, *etc.*) for its effectiveness and advancement. Through rigorous testing of this database, it was expected that the research results would not only optimize the quality control process of agricultural products and improve the efficiency of storage and transportation, but also positively affect the testing technology of the agroindustry and push the industry towards a smarter and more efficient monitoring and analysis system.

Materials and methods

Principles of the algorithm

The specific structure of the modules was shown in Figure 1. The data preprocessing module was to crop, scale, normalize, and other operations on the images of agricultural products to adapt to the input requirements of the neural network, as well as to increase the diversity and robustness of the data. The cropping step included to find a smallest rectangular region R based on the edge information of each produce image I. So that, R contained the main part of the produce, and then crop I to R to get the cropped image I_c [11]. The scaling step was that, for each cropped image I_c, according to the input size of the neural network, I_c was scaled equally to the size of $W \times H$ to get the scaled image I_s, where W and H were the pre-set width and height [12]. The normalization step was that, for each scaled image I_s, according to the distribution of its pixel values, each pixel value x of I_s was converted to z to get the normalized image I_n, where the formula for z calculation was shown in Equation (1) [13].

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

where μ and σ were the mean and standard deviation of the pixel values of I_s, respectively. The last step for data enhancement included randomly applying some image transformations such as mirroring, rotating, translating, distorting, filtering, contrast adjustment, *etc.* to each normalized image I_n to a certain probability to obtain the enhanced image I_n, in order to increase the diversity and robustness of the data [14].

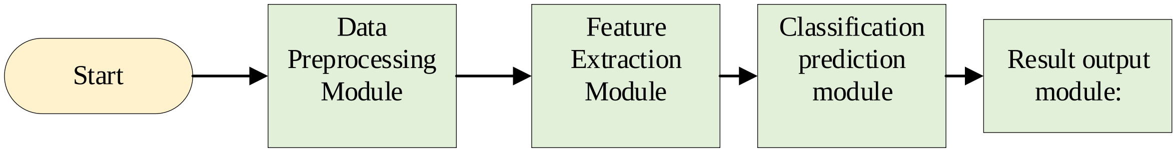

Figure 1. Module flowchart.

The feature extraction module was to use CNN and attention mechanism to extract features from the produce image, get the feature vector of the damage region, highlight the features of the damage region, suppress the features of the background and interference, and improve the differentiation and robustness of the features. For each enhanced image I_a, the feature extraction was performed on I_a using the pre-trained CNN model *F* to obtain the feature map *X*, where X had the shape of $C \times M \times N$, where *C* was the number of channels of the feature map, and *M* and *N* were the height and width of the feature map. The formula for *X* calculation was shown in Equation (2) [15].

$$X = F(I_a) \tag{2}$$

For each feature map *X*, the attention mechanism model *G* was used to weight *X* to get the attention feature map X_a, where the shape of X_a was the same as that of *X* and was calculated using Equation (3) [16].

$$X_a = G(X)\mathrm{X} \tag{3}$$

where *G(X)* was the attention weight matrix, which also had the shape of $C \times M \times N$, and the formula for *G(X)* calculation was shown in Equation (4) [17].

$$G(X) = \sigma(W_2\delta(W_1X)) \tag{4}$$

where σ was the Sigmoid activation function, δ was the relu activation function. W_1 and W_2 were the learnable parameters of the attention mechanism model, which were shaped as $rc \times C$ and $C \times rc$, respectively, with r as the scaling factor, which was generally taken as 16 or 32 [18]. The feature vector was that, for each attention feature map X_a, flatten X_a into a one-dimensional vector *x* to obtain the feature vector *x*, where *x* was of length $C \times M \times N$. The formula for *x* calculation was shown in Equation (5) [19].

$$x = flatten(X_a) \tag{5}$$

The classification prediction module was to use LSTM and MLP to classify and predict the feature vectors and get the multi-label output of the damage including the type, degree, and location of the damage, which integrated the type, state, and angle information of the agricultural products, and improved the accuracy and efficiency of the classification. For each feature vector *x*, the LSTM model *H* was utilized to process *x* to obtain the sequence output *z*, where *z* had the shape of $L \times D$, where *L* was the length of the sequence and *D* was the dimension of the sequence. The formula of z was then shown in Equation (6) [20].

$$z = H(x) \tag{6}$$

where *H* was the hidden state of LSTM and its update formula was shown in Equation (7).

$$\begin{aligned} i_t &= \sigma(W_{ii}x_t + W_{hi}h_{t-1} + b_i) \\ f_t &= \sigma(W_{if}x_t + W_{hf}h_{t-1} + b_f) \\ o_t &= \sigma(W_{io}x_t + W_{ho}h_{t-1} + b_o) \\ g_t &= \tanh(W_{ig}x_t + W_{hg}h_{t-1} + b_g) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t h_t = o_t \odot \tanh(c_t) \end{aligned} \tag{7}$$

where $i_t, f_t, o_t, g_t, c_t, h_t$ were the input gate, oblivion gate, output gate, candidate memory, cell state, and hidden state at moment t. σ was the Sigmoid activation function. $\tanh$ was the Hyperbolic Tangent Activation Function. $W_i, W_f, W_o, W_g, b_i, b_f, b_o, b_g$ were the learnable parameters of LSTM. For each sequence output *z*, the MLP model *J* was utilized for classification prediction of *z* to get the multi-label output of damage *y*, where *y* had the shape of *K*, and *K* was the number of categories of damage. The formula of *y* calculation was shown in Equation (8) [21].

$$y = J(z) \tag{8}$$

where *J* was the output layer of the MLP, which was calculated as shown in Equation (9).

$$J(z) = \sigma(W_J z + b_j) \tag{9}$$

where σ was the Sigmoid activation function, and W_J and b_j were the learnable parameters of the MLP with shapes $K \times D$ and *K*, respectively [22].

The result output module was to annotate and display the images of agricultural products according to the results of classification prediction, output the quality assessment of agricultural products and processing suggestions, and provide effective technical support and solutions for agricultural production and distribution [23]. The labeling process was to get the labeled image I_b for each enhanced image I_a and the corresponding damage of the multi-label output *y* according to the value of *y*, labeling I_a, such as marking the damage area with a red box, marking the type, degree, and location of the damage with text, *etc*. [24]. For each labeled image I_b, it would be displayed for the user to view. For each damage the multi-label output *y*, according to the value of *y*, the quality of the produce was evaluated such as giving the grade, quality, and value of the produce, *etc*. The evaluation result was obtained as r_e, which was calculated using Equation (10) [25].

$$r_e = f(y) \tag{10}$$

where *f* was an evaluation function that gave different evaluation scores and grades, such as A, B, C, *etc*., according to different types, degrees, and locations of damage. For each damage, the multi-label output *y* provided suggestions for the handling of the produce based on the value of *y*, such as giving precautions and methods for storage, transportation, processing, and consumption of the produce. The obtained suggestion result r_s was calculated using Equation (11) [26].

$$r = g(y) \tag{11}$$

where g was a suggestion function that gave different suggestions on what and how to do, such as refrigerate, excise, cook, *etc*., depending on the type, extent, and location of the damage [27]. The final output would splice the evaluation result r_e and the recommendation result r_s into a string r and output r to the user for reference.

Algorithmic process

The algorithmic process flowchart was shown in Figure 2. Using CNN and attention mechanism to construct a feature extraction network consisted of multiple convolutional layers, pooling layer, activation layer, and attention layer, which could automatically learn the feature representation of the damage on the surface of the agricultural products, highlight the features of the damaged area, suppress the features of the background and interference, and improve the differentiation and robustness of the features. The training set images were input into the feature extraction network, and the parameters of the network were optimized by the back propagation algorithm to obtain the feature vector of the training set images (Figure 3) [28]. LSTM and MLP were used to construct a classifier network, which consisted of one LSTM layer and one MLP layer, and could comprehensively consider the type, state, and angle information

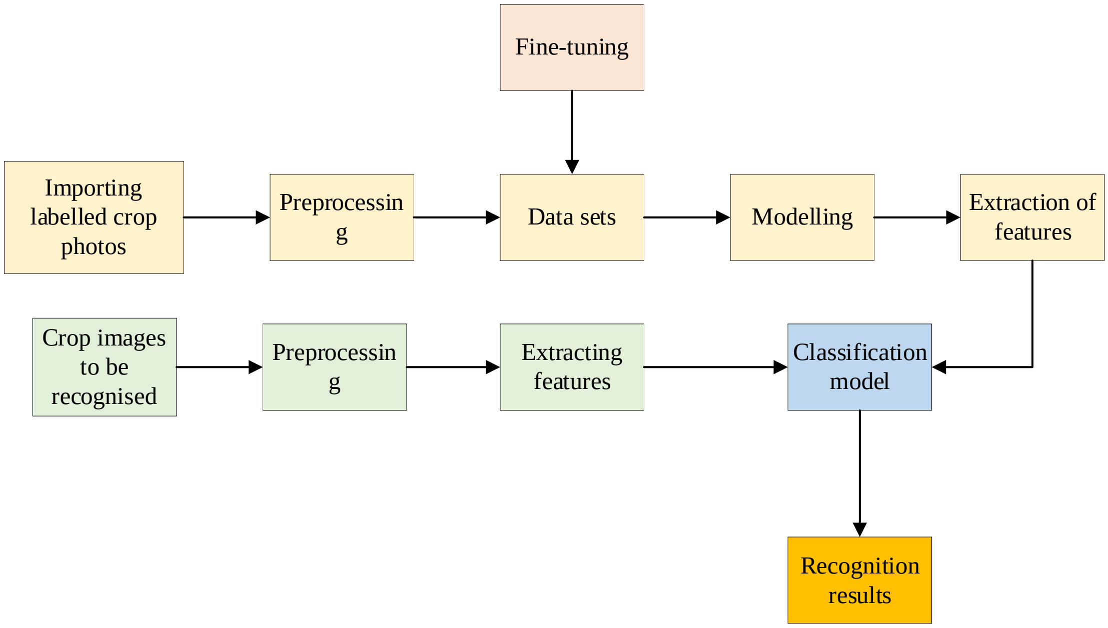

Figure 2. Algorithm flowchart.

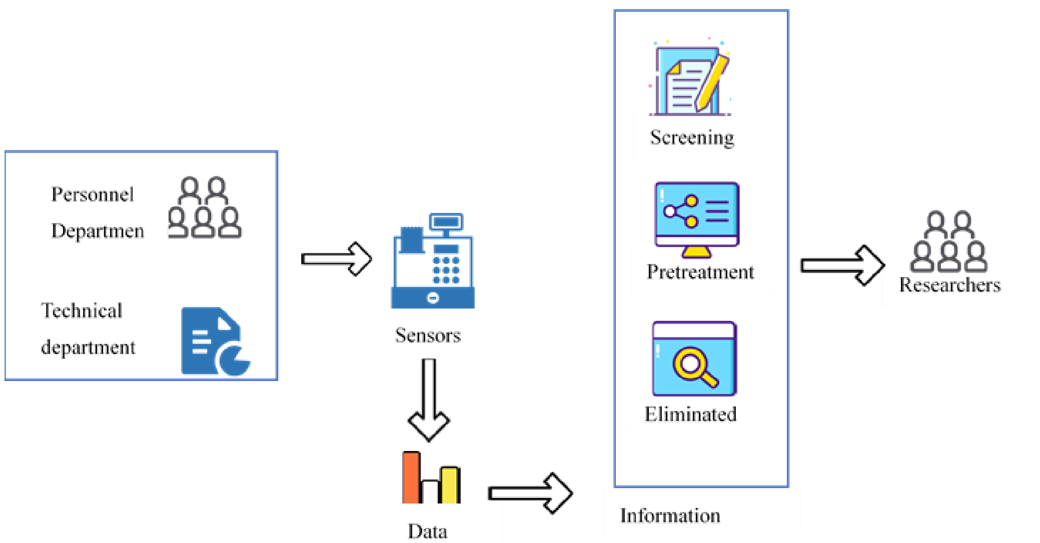

Figure 3. Data processing model.

of agricultural products to classify the surface damage of agricultural products with multiple labels, and realize the simultaneous judgment of the type, degree, and position of damage to improve the accuracy and efficiency of classification. The feature vectors and corresponding labels of the images in the training set were input into the classifier network, and the parameters of the network were optimized by the back-propagation algorithm to obtain the model of the classifier network. The role of the LSTM layer was to take advantage of the characteristics of LSTM network, and take into account the type, state, and angle information of the agricultural products comprehensively, and to convert the feature vectors of the output of the feature extraction network into a fixed-length vector, which was used as the input to the MLP layer. The LSTM layer was used as the input to the MLP layer. The parameters of the LSTM layer included the input dimension, output dimension, weights and biases of the gating units, *etc*. The role of the MLP layer was to utilize the characteristics of the MLP to perform a nonlinear transformation on the vector output from the LSTM layer, and to output a multidimensional vector with each dimension corresponding to the label of a damage, which indicated the presence or absence of the damage, the degree, and the location of the damage, *etc*. The parameters of the MLP layer included the number of hidden layers, the degree of the damage, and the location of the damage. Parameters included the number of hidden layers, the number of neurons, the activation function, the weights, and the bias [29-32].

Data sets

PlantVillage (https://github.com/PlantVillage) data used in this study was a publicly available database of plant leaf images containing 30 agricultural products and 30 injuries. The proportion of images in each category in the training set and test set was the same with 80% of the total number of images in a certain category being used as the training set and 20% of the total number of images in the same category being used as the test set. Various produces including pomegranates, tomatoes, corn, grapes, strawberries, *etc*. and different diseases were included in the data sets [33].

Assessment of the proposed model

To make a comprehensive assessment of the effectiveness of the model, the following mainstream algorithms for automatic identification and classification of surface damage on agricultural products were selected as comparison methods, which included six deep learning-based algorithms (SVM, CNN, DBN, CAE, Resnet, ACNN) for automatic identification and classification of surface damage in agricultural products. All those methods have evolved gradually from 2006 to 2022 through traditional feature extraction and classification methods to the use of techniques, which include convolutional neural networks, deep belief networks, convolutional selfencoders, deep residual networks, and attentional mechanisms to improve the performance and efficiency of identification and classification [34].

Results and discussion

Assessment of proposed model

The amount of pre-training data, the running time of the model, and the generalization ability of the model were selected in this study as the assessment experimental indicators. Four metrics including accuracy, recall, precision, and F1-score were used to determine the generalization ability of the models [35]. The results showed that the proposed model (LSTM + MLP + CNN) was the best in all the evaluation metrics, indicating that it had strong time series prediction ability, which might be due to the fact that it combined the long-term memory capability of LSTM, the nonlinear mapping capability of MLP, and the local feature extraction capability of CNN, which enabled it to better capture the dynamics and complex patterns of time series. In contrast, the other models demonstrated weak points in certain metrics such as lower recall for SVM, longer running time for CNN, higher amount of pre-training data for DBN, lower precision for CAE, lower F1 value for Resnet, and longer running time for ACNN (Table 1) [36].

Relationship between pre-training data size and the model accuracy

To explore the relationship between the amount of pre-training data and the model accuracy for the LSTM + MLP + CNN model of this study, a quiz in the same experimental setting was conducted. The results showed that the model reached the highest accuracy rate when the pre-training data reached 40,000 (Figure 4).

Table 1. Comparison of proposed LSTM + MLP + CNN model with other six deep learning-based algorithms.

Model	Number of pre-training data	Running time	Accuracy	Recall rate	Precision	F1 value
SVM	1,000	10 s	0.75	0.72	0.77	0.74
CNN	20,000	20 s	0.82	0.79	0.84	0.81
DBN	30,000	30 s	0.86	0.83	0.88	0.85
CAE	40,000	40 s	0.89	0.87	0.91	0.89
Resnet	60,000	50 s	0.92	0.90	0.93	0.91
ACNN	60,000	60 s	0.95	0.94	0.96	0.95
LSTM + MLP + CNN	40,000	70 s	0.99	0.99	0.99	0.99

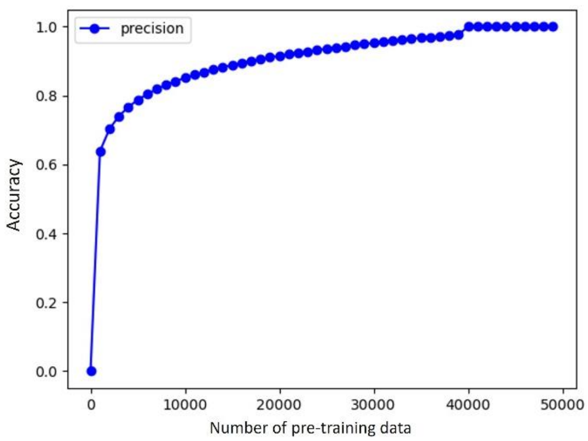

Figure 4. Relationship between model accuracy and model pre-training data.

This study proposed a deep learning-based algorithm for automatic identification and classification of damage on the surface of agricultural products, which utilized the advantages of CNN, attention mechanism, LSTM, and MLP, and was able to effectively extract the features of the surface of agricultural products and determined whether there was damage, as well as the type and degree of damage. The model combined a variety of deep learning techniques, fully utilized their advantages, and improved the performance and efficiency of the model. The evaluation indexes for this proposed model were selected comprehensively, and the evaluation results were objectively analyzed, demonstrating the superiority and stability of the model. However, although the proposed model in this study achieved the highest accuracy rate when the pre-training data reached 40,000, it did not indicate whether the selection of this data amount was general and representative. In addition, the performance and changes of the proposed model were not explored under different data amounts, therefore, data sensitivity and robustness of the model should be examined in future study.

Acknowledgements

This study was supported by Key Scientific Research Project Plan of Colleges and Universities in Henan Province (Grant No. 23A520056).

References

1. Bai YB, Gao C, Singh S, Koch M, Adriano B, Mas E, *et al*. 2018. A framework of rapid regional tsunami damage recognition from post-event TerraSAR-X imagery using deep neural networks. IEEE Geosci Remote Sens Lett. 15(1):43-47.
2. Bai ZL, Liu TJ, Zou DJ, Zhang M, Zhou A, Li Y. 2023. Image-based reinforced concrete component mechanical damage recognition and structural safety rapid assessment using deep learning with frequency information. Autom Constr. 150:19.
3. Basati Z, Jamshidi B, Rasekh M, Abbaspour-Gilandeh Y. 2018. Detection of sunn pest-damaged wheat samples using visible/near-infrared spectroscopy based on pattern recognition. Spectrochim Acta A Mol Biomol Spectrosc. 203:308-314.
4. Cai M, Cheng F, Jiang ZZ. 2023. Online tip damage diagnosis of atomic force microscope based on statistical pattern recognition. J Vib Eng Technol. 25 August 2023. https://doi.org/10.1007/s42417-023-01111-3
5. Dang XC, Shang X, Hao ZJ, Su L. 2022. Collaborative road damage classification and recognition based on edge computing. Electronics. 11(20):18.

6. Fu J, Yuan HK, Zhao RQ, Chen Z, Ren, LQ. 2020. Peeling damage recognition method for corn ear harvest using RGB image. Applied Sciences-Basel. 10(10):15.
7. Ha CT, Tageldein MM, Harding SM. 2024. The entanglement of DNA damage and pattern recognition receptor signaling. DNA Repair. 133:8.
8. He R, Zhu YF, He W, Chen H. 2018. Structural damage recognition based on perturbations of curvature mode shape and frequency. Acta Mech Solida Sin. 31(6):794-803.
9. Deng SN, Wu D, Yang JS, Luo H, Fu LL, Schmidt R, *et al*. 2022. Damage recognition of glass fiber composite bi-directional corrugated sandwich cylindrical panels via non-contacted vibration method. Mater Today Commun. 32:15.
10. Du SQ, Yang ZK, Qin YG, Wang SS, Duan HX, Yang XL. 2018. Computational investigation of the molecular conformation-dependent binding mode of (E)-β-farnesene analogs with a heterocycle to aphid odorant-binding proteins. J Mol Model 24:70.
11. Hong XB, Zhou JX, Huang GJ, Ni L. 2018. Synergetic damage recognition approach for messenger wire in icing environment using piezoceramic transducers. Measurement. 122:522-531.
12. Jin ZH, Teng S, Zhang JQ, Chen GF, Cui FS. 2022. Structural damage recognition based on filtered feature selection and convolutional neural network. Int J Struct Stab Dyn. 22(12):23.
13. Kusakabe M, Onishi Y, Tada H, Kurihara F, Kusao K, Furukawa M, *et al*. 2019. Mechanism and regulation of DNA damage recognition in nucleotide excision repair. Genes Environ. 41:6.
14. Kuznetsov NA, Fedorova OS. 2020. Kinetic milestones of damage recognition by DNA glycosylases of the helix-hairpin-helix structural superfamily. In D. O. Zharkov (Ed.), Mechanisms of Genome Protection and Repair (Vol. 1241, pp. 1-18).
15. Liu MT, Wang L, Qin ZN, Liu JQ, Chen J, Liu XM. 2020. Multi-scale feature extraction and recognition of slope damage in high fill channel based on Gabor-SVM method. J Intell Fuzzy Syst. 38(4):4237-4246.
16. Lu G, Li S, Zhou CW, Qian X, Xiang Q, Yang TQ, *et al*. 2019. Tenuivirus utilizes its glycoprotein as a helper component to overcome insect midgut barriers for its circulative and propagative transmission. PLOS Pathog. 15(3):29.
17. Lu T, Dooms A. 2021. Bayesian damage recognition in document images based on a joint global and local homogeneity model. Pattern Recognit. 118:15.
18. Modarres C, Astorga N, Droguett EL, Meruane V. 2018. Convolutional neural networks for automated damage recognition and damage type identification. Struct Control Health Monit. 25(10):17.
19. Mu H, Geacintov NE, Broyde S, Yeo JE, Schärer OD. 2018. Molecular basis for damage recognition and verification by XPC-RAD23B and TFIIH in nucleotide excision repair. DNA Repair. 71:33-42.
20. Pan TB, Zheng YL, Zhou YB, Luo WB, Xu XB, Hou CY, *et al*. 2023. Damage pattern recognition for corroded beams strengthened by CFRP anchorage system based on acoustic emission techniques. Constr Build Mater. 406:15.
21. Petrikovszki R, Cseresnyés I, Bárányos F, Molnár AG, Boros G. 2023. Early detection of root-knot nematode (*Meloidogyne incognita*) infection by monitoring root dielectric response non-destructively. Int Agrophys. 37(2):179-187.
22. Riseh RS, Hassanisaadi M, Vatankhah M, Babaki SA, Barka EA. 2022. Chitosan as a potential natural compound to manage plant diseases. Int J Biol Macromol. 220:998-1009.
23. Salehi H, Chakrabartty S, Biswas S, Burgueño R. 2019. Localized damage identification in plate-like structures using self-powered sensor data: A pattern recognition strategy. Measurement. 135:23-38.
24. Shan QL, Xu QK, Xue YD, Lian JW, Wang YW, Chen CR, *et al*. 2021. The tensile damage behavior of SiCf/SiC-B4C after oxidation in wet atmosphere based on AE pattern recognition. J Am Ceram Soc. 104(8):4131-4144.
25. Shi LF, Zhang F, Xia JS, Xie JB. 2023. Scene-level buildings damage recognition based on cross conv-transformer. Int J Digit Earth. 16(2):3987-4007.
26. Shin HK, Ahn YH, Lee SH, Kim HY. 2020. Automatic Concrete Damage Recognition Using Multi-Level Attention Convolutional Neural Network. Materials. 13(23):13.
27. Shu JP, Ding W, Zhang JW, Lin FZ, Duan YF. 2022. Continual-learning-based framework for structural damage recognition. Struct Control Health Monit. 29(11):17.
28. Silik A, Noori M, Ghiasi R, Wang TY, Kuok SC, Farhan NSD, *et al*. 2023. Dynamic wavelet neural network model for damage features extraction and patterns recognition. J Civ Struct Health Monit. 13(4-5):925-945.
29. Thirumalaiselvi A, Sasmal S. 2021. Pattern recognition enabled acoustic emission signatures for crack characterization during damage progression in large concrete structures. Appl Acoust. 175:12.
30. Tu JC, Yu LC, Zhao J, Zhang JZ, Zhao D. 2022. Damage modes recognition of wood based on acoustic emission technique and hilbert-huang transform analysis. Forests. 13(4):22.
31. Van der Zee YJ, Stiers PLJ, Evenhuis HM. 2022. Object recognition and dorsal stream vulnerabilities in children with early brain damage. Front Hum Neurosci. 16:13.
32. Wang F, Wang C, He F. 2021. Simulation of recognition method of damaged parts of high-intensity sports injury images. Rev Bras de Medicina do Esporte. 27(5):504-508.
33. Wu YG, Shen X, Li DH. 2021. Numerical and experimental research on damage shape recognition of aluminum alloy plate based on Lamb wave. J Intell Mater Syst Struct. 32(18-19):2273-2288.
34. Xiong FG, Wen HJ, Zhang C, Song CH, Zhou XZ. 2022. Semantic segmentation recognition model for tornado-induced building damage based on satellite images. J Build Eng. 61:14.
35. Xu J, Wang WX, Han QH, Liu X. 2020. Damage pattern recognition and damage evolution analysis of unidirectional CFRP tendons under tensile loading using acoustic emission technology. Compos Struct. 238:10.
36. Yan JF, Xie Y, Zhang QJ, Gan L, Wang F, Li HY, *et al*. 2019. Dynamic recognition and repair of DNA complex damage. J Cell Physiol. 234(8):13014-13020.