

Article

# An Image Identification System for Agricultural Items Based on Machine Vision

Meng Zhang<sup>1</sup>, Zheng Liang<sup>1,\*</sup>, and Shuai Tong<sup>2</sup>

<sup>1</sup> Social Sciences Department of Shandong Medical College, Jinan, 250002, China

<sup>2</sup> Department of Jinan Technology School, Jinan, 250001, China

\* **Correspondence:** SDMCzhangmeng@163.com

**Abstract:** The purpose of this work was to use machine learning classification models and hyperspectral camera technologies to create a model of surface damage to garlic. 140 of the 184 garlic plants of which 44 were used for test validation were pre-treated for surface damage. First, we examined the data in ENVI under various damage scenarios using the normalised vegetation index (NDVI) approach. 579 pixels were then chosen for the training of the logistic regression model. Finally, we used 54 garlic bulbs to practically validate the model. Although tiny regions could not be precisely identified, the mouldy portion of the garlic's surface could be identified using the NDVI technique. 90% accuracy was attained using the 90% classification model constructed using the logistic regression approach. Garlic's surface damage, even at first mild ones, was correctly identified. The creation of this model for identifying garlic damage lowers the cost of detecting garlic damage and broadens the use of hyperspectral technologies in garlic detection.

**Keywords:** Massive garlic superficial damage, Hyperspectral camera, Logistic regression, Machine learning, NDVI

---

## 1. Introduction

One of the biggest issues facing the agricultural sector in recent years has been the issue of pests and illnesses in crop production [1]. Surface damage is widespread in garlic farming, but it can be challenging to monitor and identify properly. The severity of the damage has a direct impact on yield and quality [2]. In addition to being labour- and time-intensive, traditional manual inspection techniques are also extremely susceptible to subjective factors, which makes it challenging to satisfy the needs of large-scale manufacturing [3]. To increase the productivity and calibre of garlic production, it is crucial to design an automated damage identification approach that is both accurate and efficient.

Compared to RGB and multispectral photos, high-spectrum images provide more wavelength information, which makes them valuable in a range of industries. They are also employed in security, gender studies, and agriculture, including plant classification [4]. In addition to crop pests and illnesses, hyperspectral images of fruits are utilised to identify mechanical damage [5]. It can be applied to early egg analysis and even to evaluate the quality of the eggs. There have been reports that it is mobile [6]. Research on machine learning for the selection of damaged garlic has primarily made scapular RGB cameras possible. However, hyperspectral camera technology primarily influences allicin and the soluble solids content (SSC: SolubleSolid Content; henceforth referred to as SSC) of

garlic, in addition to the garlic's sterility following air-drying. Singularity has been detected using it [7]. In order to monitor allicin and damage in garlic, simultaneous usage of RGB and hyperspectral cameras raises the expense if such damage is found. The hyperspectral camera can also identify plant damage, and research has been done to find a way to get more useful information from a single hyperspectral image while also lowering the expense of identifying garlic damage based on the optical image [8].

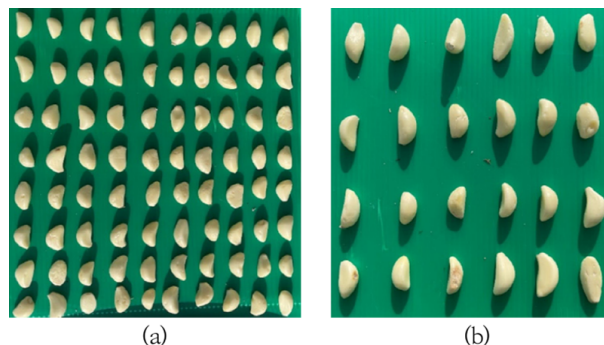
Our goal in this study was to use hyperspectral camera technology and machine learning techniques to create an efficient model for recognising damage on the surface of garlic. Out of the 154 garlic plants that we gathered data from, 140 were utilised to create the model and 44 were used to validate the test. After preprocessing the data using the normalised vegetation index (NDVI) approach, we trained a logistic regression model on 579 pixels. Lastly, in order to evaluate the precision and dependability of the model, we field validated 54 garlic bulbs [9, 10].

High-level logistic regression was used to evaluate the created model using a garlic classification model. In the end, the model was identified by comparing it with RGB photographs and using high-spectral photos of 24 garlic plants.

## 2. The Materials and Methods

### 2.1. Experimental Hardware and Data Preprocessing

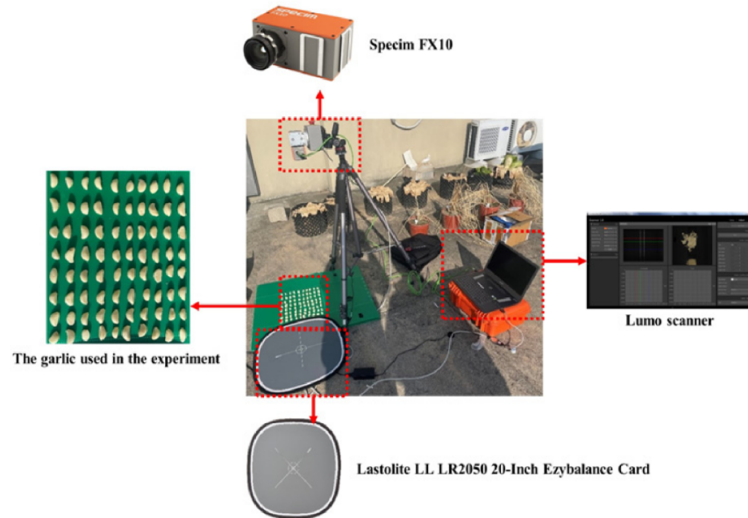
184 garlic plants were ready for use in this study after being acquired from a nearby store in Jinan, China. Using a spatula, the garlic was sliced into several tiny and big damaged areas on its surface [11, 12]. The garlic was then exposed to varying degrees of damage for approximately five days at a temperature of 25°C indoors. The decaying garlic looked like what Fig. The acquired samples are displayed in Figure 1. As seen in Figure 1, the 184 garlic plants were split into two groups. Figure 1a shows the classification model developed using eighty species of garlic; Figure 1b shows the classification model developed using the remaining species. Figure 1b illustrates how 24 garlies were finished.



**Figure 1.** Samples of Garlic Used in the Experiment: (a) Experimental Group; (b) Validation Group

A Specim FX10 was employed by the camera to capture macro-image data. This device may be used to gather spectral data and has both visible and near-infrared (VNIR) wavelengths (400-1000 nm, bands 1-56) Table 1. The Lastolite LLL LR2050 20-inch Ezybalance card 18% (Manfrotto, Italy) is the reflector that was utilised to adjust the reflectance of the photos that were obtained for the high-spectral camera [13, 14]. The desktop programme Lumo Scanner was used to capture hyperspectral data.

The garlic configuration illustrated in Figure 2 can be used to set up and regulate the canning stage and sample position for the study. To do this, place the sample roughly 1.2 metres away from the hyperspectral camera for image data gathering. Position the high-frequency camera on the same board as shown in Figure 2 and fasten it to the tripod support.



**Figure 2.** Composition of Outdoor High-Resolution Spectral Image Acquisition System

A laptop computer was linked to the hyperspectral camera’s data port, and a Lumo scanner was utilised to measure dark current images in order to reduce self-noise when the sky was clear [15]. Store the data after setting the binning value to twice the default value. Following the acquisition of the raw hyperspectral data, further modelling and computations were carried out. All pixel values were dispersed between 0 and 1 after reflectance correction was applied to the raw data using a reflector to aid in model training. (SPECIM FX10 specifications are shown in Table 1.)

Aticle	Vntilator
Sectral range	400-1000 nanometers
Sectral band	224
Spectral FWHM (Full Width at Half Peak)	55 nanometers
Satial sampling	1024 pixels
Frame rate	330FPS full frame 9900FPS, select 1 frequency band
FOV	38
F value	F/1.7
Camera signal-to-noise ratio (SNR)	600:1)
Camera Interface	Gige Vision g /mera link)

**Table 1.** SPECIM FX10 Specifications

### 2.2. Vegetation Index

The red and near-infrared bands on crops represent the normalised vegetation index, or NDVI [16]. Using line bands from multispectral pictures and reflectance variations in the red and near-infrared bands, the NDVI is a measure of the intensity of vegetation [17]. With the use of observed multispectral satellite photos, the NDVI can be used to assess the true state of crops.

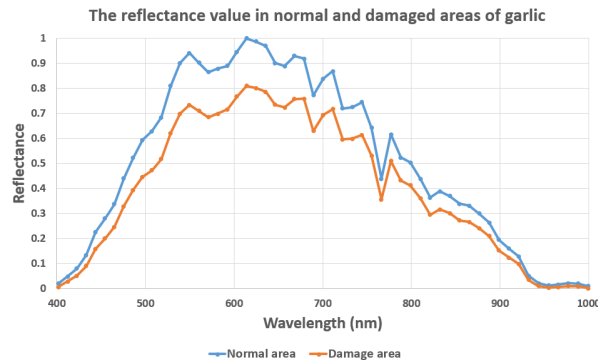
The NDVI is computed using the same formula as in Eq. (1). Here, red indicates infrared reflectance while NIR is the closest ultraviolet reflectance. NIR data are found in band 39 (810 nm), while red data are found in band 24 (650 nm). ENVI computes the NDVI picture and generates the resulting image.

$$NDVI = (NIR - Red)/(NIR + Red). \tag{1}$$

The near-infrared reflectance is denoted by NIR, while the infrared reflectance is indicated by Red.

### 2.3. Development and Validation of Classification Models

As seen in Figure 1, the ROI function was used to separate the substantially damaged and normal regions in ENVI prior to model creation. The average reflection values in each band for the injured and normal portions were calculated, same like in Figure 3. Throughout the whole band from 400 nm to 1000 nm, the reflectance values in the damaged zone are different from those in the normal region. This led to the establishment of 56 bands as the foundational data for the ensuing model training.



**Figure 3.** Shows the Reflectance Values of Both Healthy and Injured Garlic

Little computing work is needed to use the logistic regression model. Furthermore, it is quick and adaptable to a wide range of two-dimensional classification issues [18]. The logistic regression model was selected for this study in order to create a classification model, and it was obtained and verified. The ratio of damaged to normal areas varied significantly across the 80 garlic plants in the training data; therefore, 56 bands of reflections (300 pixels) and normal ROI areas (279 pixels) were chosen for a total of 579 pixel locations in the damaged ROI areas in order to close the gap.

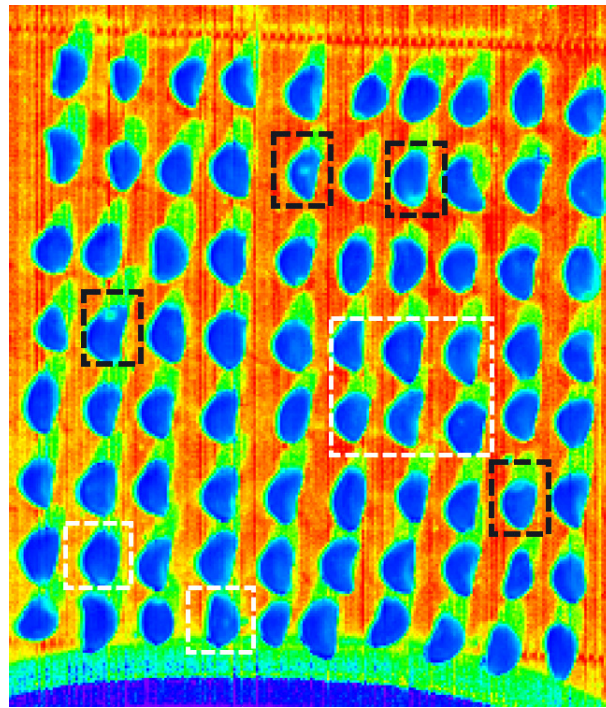
The validation team examined 24 samples in order to confirm the model's capacity for categorization. The full band's reflectance data were used to create the logic. Regression modelling was used for detection, and the ROI function in ENVI was used to examine the outcomes of the identified damaged and normal pixels. The outcomes were compared with those from the RGB camera after being shown in Jin.

### 3. Result

Figure 4 displays the typical surface area of garlic from the NDVI photo result, with the damaged portion partially visible. As can be observed, the damaged area is easily identified from the regular area, whereas the area with minimal damage in the white dashed area barely deviates from the normal area. As depicted in Fig. Garlic has deteriorated after significant injury, as evidenced by the NDVI values when compared to the RGB shot in 1a.

Table 2 displays the model's accuracy results. 0 represents normal areas, and 1 represents damaged areas. As may be shown, the model's overall identification rate is 90%. With high F1 scores of 85% and 92%, the expected failure and recall rates were 97% and 76% for normal areas and 86% and 98% for damaged areas, respectively. Figure 5 displays the overall forecast accuracy results for the model. One way to evaluate the receiver's operating characteristics is as follows: The region beneath the ROC curve becomes the area under the curve (AUC), and once more, the greater the AUC area, the higher the model's accuracy. The ROC curve is near the upper left quadrant. The AUC result for the model created in this investigation is 92%, according to the ROC curve.

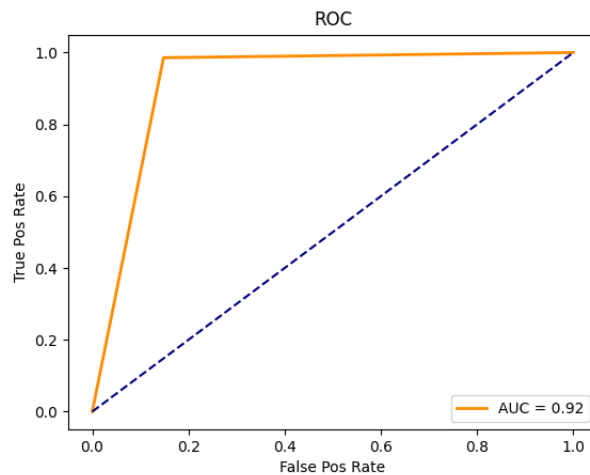
The red area in 6a is the result of anomalous damage, as seen in Figure 6. Garlic's RGB data is seen in 6b. It is evident that the logistic regression approach can correctly identify both minor and degraded areas of damaged garlic, in contrast to the identification findings of the NDVI method. Nonetheless, the model's drawbacks, as created in this study, are indicated below. In the future, we will create an



**Figure 4.** NDVI Technique Results for the Experimental Group (Black Dashed Boxes Indicate Successfully Recognised Damaged Garlic Samples, Whereas White Dashed Boxes Indicate Unsuccessfully Identified Damaged Garlic Samples)

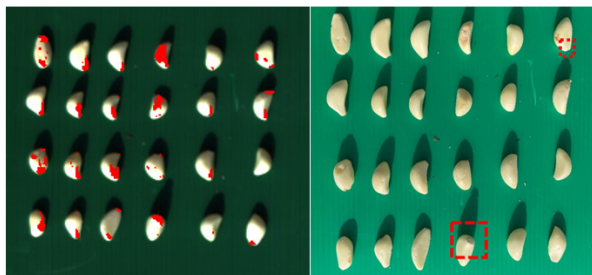
Item	Precision	Recall	F1-score
0	0.95	0.77	0.82
1	0.85	0.96	0.92
Accuracy	-	-	0.92
Macro avg	0.92	0.85	0.88
Weighted avg	0.90	0.90	0.90

**Table 2.** Regression Model Precision, Recall, and F1 Score Outcomes



**Figure 5.** ROC Results of the Regression Model)

accurate classification model that can distinguish between garlic cloves and roots, as indicated by the dashed red region in Figure 6b, which also indicates that the garlic roots were identified as damaged.



**Figure 6.** Identification Results of 24 Cloves of Garlic in the Experimental Group: (a) Model Classification Results and (b) RGB Camera Results

This work proposes a surface damage detection approach based on logistic regression that always analyses the complete waveform data. The significant cost of data processing is a drawback. Nevertheless, the conventional NDVI method uses a threshold value to assess the degree of surface damage to garlic; the threshold value's assessment may be arbitrary based on the quantity of samples and the statistical analysis technique used. Nonetheless, by providing a value of 0 or 1, the logistic regression approach suggested in this work can generate objective classification findings and is thus thought to be more trustworthy.

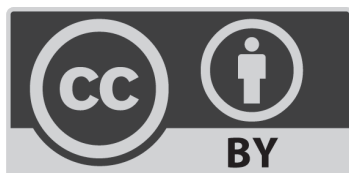
#### 4. Conclusion

This study used logistic regression modelling and high-spectral image processing techniques to create a surface damage detection model for garlic. After being preprocessed in ENVI, raw high-spectral photos of garlic outside the plant were sent back. After four corrections and normalisation processes, the NDVI image results were examined for the average reflectance values of the damaged and normal areas. A total of 579 pixels of data from the damaged and normal areas were added to the logistic regression model, which was trained and assessed for the validation group. Using this technique, the model's accuracy was confirmed. When compared to the NDVI approach, the experimental findings demonstrate that the logistic regression model's overall accuracy can exceed 90%. Additionally, the model performs well in identifying areas with both major and moderate garlic damage.

#### References

1. Ang, K. L. M. and Seng, J. K. P., 2021. Big data and machine learning with hyperspectral information in agriculture. *IEEE Access*, 9, pp.36699-36718.
2. Saha, D. and Manickavasagan, A., 2021. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review. *Current Research in Food Science*, 4, pp.28-44.
3. Nturambirwe, J. F. I. and Opara, U. L., 2020. Machine learning applications to non-destructive defect detection in horticultural products. *Biosystems engineering*, 189, pp.60-83.
4. Wu, Q., Xu, L., Zou, Z., Wang, J., Zeng, Q., Wang, Q., Zhen, J., Wang, Y., Zhao, Y. and Zhou, M., 2022. Rapid nondestructive detection of peanut varieties and peanut mildew based on hyperspectral imaging and stacked machine learning models. *Frontiers in Plant Science*, 13, p.1047479.
5. Shao, Y., Gao, C., Xuan, G., Gao, X., Chen, Y. and Hu, Z., 2020. Determination of damaged wheat kernels with hyperspectral imaging analysis. *International Journal of Agricultural and Biological Engineering*, 13(5), pp.194-198.

6. Pang, Q., Huang, W., Fan, S., Zhou, Q., Wang, Z. and Tian, X., 2022. Detection of early bruises on apples using hyperspectral imaging combining with YOLOv3 deep learning algorithm. *Journal of Food Process Engineering*, 45(2), p.e13952.
7. Fracarolli, J. A., Pavarin, F. F. A., Castro, W. and Blasco, J., 2020. Computer vision applied to food and agricultural products. *Revista Ciencia Agronomica*, 51(spe), p.e20207749.
8. Zhang, P., Shen, B., Ji, H., Wang, H., Liu, Y., Zhang, X. and Ren, C., 2022. Nondestructive prediction of mechanical parameters to apple using hyperspectral imaging by support vector machine. *Food Analytical Methods*, 15(5), pp.1397-1406.
9. Nasr-Esfahani, S., Muthukumar, V., Regentova, E. E., Taghva, K. and Trabia, M. B., 2022. Detection of Pits in Olive Using Hyperspectral Imaging Data. *IEEE Access*, 10, pp.58525-58536.
10. Zhu, M., Huang, D., Hu, X. J., Tong, W. H., Han, B. L., Tian, J. P. and Luo, H. B., 2020. Application of hyperspectral technology in detection of agricultural products and food: A Review. *Food science & nutrition*, 8(10), pp.5206-5214.
11. Liu, Y., Li, M., Wang, S., Wu, T., Jiang, W. and Liu, Z., 2020. Identification of heat damage in imported soybeans based on hyperspectral imaging technology. *Journal of the Science of Food and Agriculture*, 100(4), pp.1775-1786..
12. Chen, S. Y., Cheng, Y. C., Yang, W. L. and Wang, M. Y., 2021. Surface defect detection of wet-blue leather using hyperspectral imaging. *IEEE Access*, 9, pp.127685-127702.
13. Wang, B., Sun, J., Xia, L., Liu, J., Wang, Z., Li, P., Guo, Y. and Sun, X., 2023. The applications of hyperspectral imaging technology for agricultural products quality analysis: A review. *Food Reviews International*, 39(2), pp.1043-1062.
14. Li, X., Liu, Y., Jiang, X. and Wang, G., 2021. Supervised classification of slightly bruised peaches with respect to the time after bruising by using hyperspectral imaging technology. *Infrared Physics & Technology*, 113, p.103557.
15. Zhang, L., Rao, Z. and Ji, H., 2020. Hyperspectral imaging technology combined with multivariate data analysis to identify heat-damaged rice seeds. *Spectroscopy Letters*, 53(3), pp.207-221.
16. Zhang, H., Chen, Y., Liu, X., Huang, Y., Zhan, B. and Luo, W., 2021. Identification of common skin defects and classification of early decayed citrus using hyperspectral imaging technique. *Food Analytical Methods*, 14(6), pp.1176-1193..
17. Zhu, L., Spachos, P., Pensini, E. and Plataniotis, K. N., 2021. Deep learning and machine vision for food processing: A survey. *Current Research in Food Science*, 4, pp.233-249.
18. Yan, Y., Ren, J., Zhao, H., Windmill, J. F., Ijomah, W., De Wit, J. and Von Freeden, J., 2022. Non-destructive testing of composite fiber materials with hyperspectral imaging—Evaluative studies in the EU H2020 FibreEUUse project. *IEEE Transactions on Instrumentation and Measurement*, 71, pp.1-13.



©2024 the Author(s), licensee Combinatorial Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)