IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 March 2023, accepted 6 April 2023, date of publication 10 April 2023, date of current version 20 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3265998

RESEARCH ARTICLE

Application of X-Ray Imaging and Convolutional Neural Networks in the Prediction of Tomato Seed Viability

SUK-JU HONG[1,2,3], SEONGMIN PARK[2,4], CHANG-HYUP LEE[2,3], SUNGJAY KIM[2,3], SEUNG-WOO ROH[2,3], NANDITA IRSAULUL NURHISNA[2,4], AND GHISEOK KIM[2,3,4]

[1]Department of Agricultural Engineering, National Institute of Agricultural Sciences, Rural Development Administration, Jeonju 55364, South Korea
[2]Department of Biosystems Engineering, College of Agriculture and Life Sciences, Seoul National University, Seoul 08826, South Korea
[3]Integrated Major in Global Smart Farm, College of Agriculture and Life Sciences, Seoul National University, Seoul 08826, South Korea
[4]Research Institute of Agriculture and Life Science, College of Agriculture and Life Sciences, Seoul National University, Seoul 08826, South Korea

Corresponding author: Ghiseok Kim (ghiseok@snu.ac.kr)

This work was supported in part by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through the "Development of Advanced Agricultural Machinery Industrialization Technology Project," funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA),under Grant 320029-03-3-HD050; and in part by the "Smart Farm Innovation Technology Development Program" funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA) and the Korea Smart Farm R&D Foundation (KoSFarm) under Grant 421031042HD050.

ABSTRACT Tomato fruits are consumed worldwide owing to their health benefits, taste, and flavor. In tomato cultivation, seed viability is directly related to crop productivity. Currently, the methods used to evaluate seed viability involve destructive sampling tests; accordingly, nondestructive methods for predicting seed viability are urgently required. This study aimed to develop X-ray imagery-based models capable of predicting the viability of tomato seeds. Particularly, X-ray-imaged seeds were grown to the seedling stage, and seedlings were classified following their condition. The structural integrity of the seeds was calculated from the X-ray image processing, and an integrity-based viability prediction model was evaluated. Furthermore, convolutional neural network (CNN)-based viability prediction models were developed and evaluated. Both models showed strong performance in distinguishing germinated and non-germinated seeds. However, the CNN-based model revealed greater accuracy in seed viability prediction than the image-processing-based model. The CNN-based model accuracy was 86.01%, with an F1 score of 92.11%, indicating the usefulness of the developed nondestructive testing approach for evaluating tomato seed viability.

INDEX TERMS X-ray image, tomato, seed viability, deep learning, convolutional neural networks.

I. INTRODUCTION

Tomatoes (*Solanum lycopersicum* L.) are consumed in different forms worldwide, including curries, sauces, and salads. Tomato consumption is associated with health benefits owing to its anti-inflammatory, antigenotoxic, antimutagenic, and antiproliferative properties [1], [2], [3], [4], [5]. In particular, lycopene, which is found at high levels in tomatoes, is a highly effective antioxidant known to reduce the risk of certain cancer types [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Ravibabu Mulaveesala.

Seed quality directly affects crop yield and is vital for optimizing crop production costs [7]. Seed viability is a critical factor affecting seed quality, which varies following genetic variation and can deteriorate during seed storage and distribution. Several methods, including standard germination, electrical conductivity, accelerated aging, and tetrazolium tests, have been proposed for assessing seed viability [8], [9], [10]. However, these methods have several drawbacks, including being invasive and labor intensive and having long testing periods [10]. Additionally, only a sampling test is possible through these destructive methods, and sorting low-quality seeds remains impossible. Therefore, an automated method

of sorting seeds using nondestructive viability assessment technology is highly necessary.

Nondestructive testing of agricultural products is a growing area of interest in the agricultural industry, and research is being conducted using various nondestructive measurement technologies. In particular, X-ray imaging technology has been used to evaluate the internal structures of various agricultural products. Notably, it has been applied for fish bone detection [11], tenderness measurement for meat [12], detection of seed infection [13], internal defect detection in nuts [14], [15], [16], [17], and fruits [18], [19], [20], [21], fruit water content measurement [22], and fruit microstructure measurement [23].

The internal structures of seeds, including the endosperm and embryo, are associated with seed viability [24]. Accordingly, X-ray images have been used to assess the internal structures of seeds for viability prediction. For example, Burg et al. [25] predicted the morphology of seedlings based on the X-ray analysis of tomato seeds, Silva et al. [26] studied the association between X-ray imaging and tomato seed germination, and Gomes-Junior et al. [27] analyzed seed density using X-ray images to assess melon and watermelon seed quality. Notably, these previous studies' findings revealed that seeds' internal cavity space is closely related to seed viability.

Prediction modeling is required for quantitative predictions from X-ray images. Conventional image-based modeling approaches are based on feature-extraction methods, including pixel intensity [28], [29], [30] and edge features [14], [20], [31]. Recently, deep neural networks that learn appropriate features from data have been widely used in vision research. Convolutional neural networks (CNNs) have been state-of-the-art in image classification tasks for many years since Alexnet [32] took first place at ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. After the proposal of the vision transformer [33] in the image classification on ImageNet, models using a vision transformer or a combination of CNN and a vision transformer show higher performances than CNN models. However, because of the low inductive bias of the vision transformer, CNNs are still widely applied to vision tasks with low complexity and a limited amount of data. For example, Ahmed et al. [34] compared traditional machine learning and a CNN-based model for watermelon seed viability assessment, and Medeiros et al. [35] used X-ray images and a CNN to predict the seed vigor of the oilseed crop *Crambe abyssinica*.

Several previous studies reported the relationship between the internal structure and viability of seeds, but most of them analyzed only the qualitative relationship. In particular, in studies analyzing the relationship between X-ray images and the viability of tomato seeds, internal structures such as cotyledon configuration or internal cavity area were investigated without prediction modeling [25], [26]. There are several factors to be considered in the internal structure of tomato seeds, such as cavity area, embryo configuration, and damage. Additionally, tomato seed embryos can exhibit diverse forms, such as 'coiled' and 'spectacles' structures [25]. Considering the intricate internal structure of tomato seeds, it is imperative to develop prediction models that can account for multiple internal features to achieve accurate predictions of seed viability.

The development of normal and vigorous seedlings is an important factor in evaluating seed quality [36]. In fact, even if the seeds germinate, seedlings that are not vigorous or have defects are not commercially viable. However, several nondestructive seed viability prediction studies have been conducted based only on seed germination, often using paper-based germination tests. Since seed viability can be assessed using a range of criteria, the performance of prediction methods can be assessed using multiple criteria. Therefore, it is necessary to analyze which viability criterion is most applicable to the applied nondestructive testing technology.

This study aimed to evaluate tomato seeds' quality by assessing the seed's internal morphology using X-ray images. X-ray images were obtained for tomato seeds in their natural state, which had not been artificially aged, and seed viability was tested using multiple criteria. The image processing method was designed to quantify the internal integrity of seeds, and an integrity-based viability prediction model was developed. Furthermore, CNN models were trained and evaluated, and their performances were compared with that of the integrity-based model. Through these processes, our study attempted to quantitatively analyze the relationship between the internal cavity area and the viability of tomato seeds, which had been qualitatively studied in previous studies. Furthermore, we aimed to improve the accuracy of predicting seed viability by developing CNN-based models that can use multiple features other than the cavity area.

II. X-RAY IMAGING AND SEED VIABILITY TEST

A. SAMPLE PREPARATION

In total, 1,152 tomato seeds of the "TY sweetiny" cultivar and 576 seeds of the "Tiniup" cultivar were used for X-ray imaging and viability test experiments. According to a previous study [24], the internal structure of seeds related to viability became clear after seed priming. Therefore, the tomato seeds in this study were used after seed priming.

B. X-RAY IMAGE ACQUISITION

An X-ray imaging device (X-eye SF160NCT, SEC Co. Ltd., Suwon, Republic of Korea) was used for image acquisition. The tube voltage and current of the X-ray imaging device were set at 70 kV and 300 μA, respectively. The prepared tomato seeds were placed in an 8 $\times$ 8 square-well plate, as shown in Fig. 1(a). The size of one well was 7 mm $\times$ 7 mm, and the plate material was acrylic with high X-ray transmittance to prevent interference during x-ray imaging. Images were acquired at a resolution of 1,280 $\times$ 1,280 pixels, and the field of view was set to 20 mm $\times$ 20 mm so that four

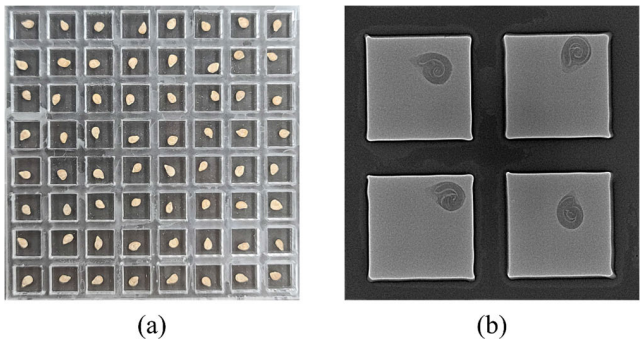

FIGURE 1. (a) Seed samples prepared on an acrylic plate; (b) an acquired X-ray image of tomato seeds.

wells could be measured in one image. Therefore, four tomato seeds were captured in one image, as shown in Fig. 1(b).

C. VIABILITY TEST

After X-ray imaging, the prepared tomato seed samples were sown in the soil in seedling trays. The samples were grown for two weeks in a greenhouse at 25–30°C, and the condition of the seedlings was classified into the following four classes by quality inspectors working at a seed company: first-grade (with a size >70%), second-grade (with a size of 50–70%), abnormal (with a size of <50% or having defects), and non-germinated (when germination did not occur). Percentages of size were calculated based on the size of the well-grown seedling, considered to be of the highest quality. In seed quality studies, the viability evaluation criteria are often divided into normal seedling, abnormal seedling, and non-germinated seed [37], [38], [39], [40]. However, there can be differences in the degree of vigor among normal seedlings. Therefore, normal seedling was also divided into first-grade and second-grade to verify the possibility of distinguishing vigor through X-ray imaging. The criteria for classifying the viability class followed the criteria used for product evaluation by Pan Pacific Seed (PPS), a seed company in the Republic of Korea.

III. MODELING FOR VIABILITY PREDICTION

A. INTEGRITY QUANTIFICATION

Previous studies have reported that the structural integrity of a seed can be associated with seed viability [26], [27], [35], [37]. Seed with a smaller cavity area has a higher amount of space occupied by the embryo and endosperm, indicating high germination potential. Therefore, to quantify the integrity of the seeds and compare it with the viability test result, an image-processing algorithm was developed to quantify the internal cavity space of the seeds using the X-ray images. First, a Gaussian blur method was employed to reduce errors during binarization. Each image was inversely binarized using the Otsu threshold, and a closing transformation was applied to the binarized image to reduce noise. Subsequently, the flood fill method was applied to fill the internal area of the seed, and erosion transformation was used

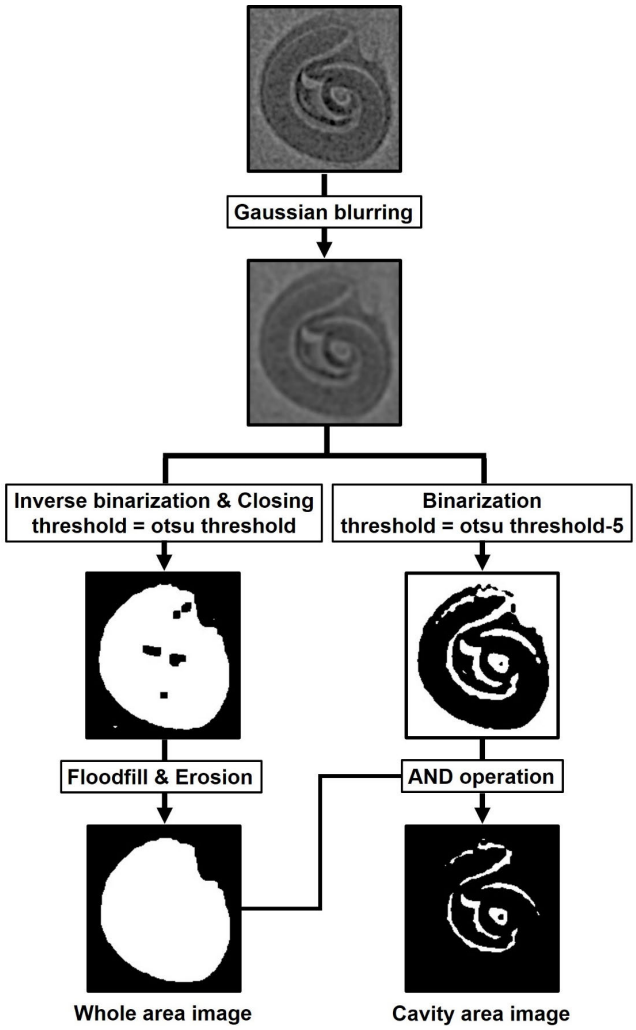

FIGURE 2. Image-processing workflow for the quantification of seed integrity.

to remove the boundary of the foreground object to determine the area of the whole seed. For extraction of the whole area of the seed, the Gaussian blurred images were binarized with the Otsu threshold minus 5. Next, the internal area images of the seeds were derived using an "AND" operation for the binarized images and the whole seed area image. Finally, the integrity of the seeds was calculated using (1). A visualization of the image processing procedure for seed integrity quantification is shown in Fig. 2.

$$\text{Seed integrity (\%)} = \frac{1 - \text{cavity area of seed}}{\text{whole area of seed}} \times 100 \quad (1)$$

B. CNN MODELING

CNN models were customized and optimized using two base models, ResNet [41] and VGGNet [42], which are the representative CNN structures frequently used in vision studies. CNN models have different model sizes depending on their structure, and some models have deep and large structures for complex features. However, the benefits of training large-size models on small datasets are limited [43]. In this study, the

optimization results showed that smaller-sized models tended to perform better that the larger-sized models. Therefore, the adopted optimized models had fewer layers than the original models.

Each convolution layer in the VGGNet architecture has a 3×3 convolution kernel. VGGNet can construct deeper layers with small parameters without excessive image size reduction due to its small convolution kernel size. Our VGGNet-based model adopted this 3×3 convolution kernel with a 1×1 stride in all convolution operations. A 2×2 maximum pooling layer was added to each of the two convolution layers. Additionally, one dense layer, with 512 nodes, was used (Fig. 3(a)).

The deep layers of a neural network enrich the feature levels, and as such, the depth of a network is considered a crucial factor for improving model performance. When a CNN layer becomes deeper, gradient vanishing and exploding issues may occur. Accordingly, the training of deep convolutional networks becomes challenging. Therefore, a deep plain network induces increased training and testing errors. To overcome these issues, He et al. [41] suggested adding skip connections, whereby the input skips layers and is added to the output of that layer or layers. Residual learning from such skip connections reduces the degradation of deep networks, leading to the effective training of deeper models. A key feature of ResNet is its bottleneck architecture, which uses 1×1, 3×3, and 1×1 convolution layers instead of two 3×3 convolution layers to reduce the amount of computation. In this study, a ResNet-based model was optimized based on skip connections and bottleneck architectures, as illustrated in Fig. 3(b).

C. DATASET SEGREGATION

X-ray images and seedling class data were divided into training, validation, and test sets at a ratio of 3:1:1. Of the 1,723 data points, 1,037 were included in the training set, and 343 were included in the validation and test sets. The training set was used for model training, and the validation set was used for model optimization. The evaluation metrics of the model were calculated using the test set.

D. MODEL TRAINING AND EVALUATION

In our study, binary classification models were trained and evaluated on multiple viability criteria. In the viability test, the seeds were classified into four classes (first-grade, second-grade, abnormal, and non-geminated). The criteria for classifying the first-grade class from the other classes were named "cut1", the criteria for classifying the first- and second-grade classes from the other classes were named "cut2", and the criteria for classifying the non-germinated class from the other classes were named "cut3". These three criteria were based on seed companies' standards for determining the seed lot quality. Table 1 summarizes the binary classification criteria. The CNN-based binary classification models trained for the cut1, cut2, and cut3 criteria were also evaluated for each test criterion of cut1, cut2, and cut3. It is not appropriate to use a 50% confidence classification threshold as the optimal threshold when evaluating the usage of criteria other than the trained criteria. Therefore, the area under the curve of the receiver operating characteristic (ROAUC), which is independent of the confidence score threshold, was used as an evaluation metric. The cut1, cut2, and cut3 models in this study refer to models trained and optimized with training and validation datasets based on cut1, cut2, and cut3 criteria, respectively. Similarly, cut1-, cut2-, and cut3-tested results mean that the results were evaluated on the test dataset classified based on cut1, cut2, and cut3 criteria. Horizontal/vertical flip, rotation, brightness adjustment, and zoom in/out were used as the data augmentation methods.

The class imbalance problem can affect model training and evaluation. Here, because the seeds were not artificially degraded, the number of seeds in the non-germinated and abnormal classes seeds was lower than in the other classes. The class balance was achieved via data oversampling during the training, and weighted metrics were calculated during the evaluation to prevent bias towards the majority class during training and evaluation. The imbalance ratio (IR) was calculated using Equation (2) and used as balancing weight. Equations (3)–(11) describe the evaluation metrics used in this study, where TP is the number of true positive samples, FP is the number of false positive samples, TN is the number of true negative samples, and FN is the number of false negative samples.

$$\text{Imbalance ratio (IR)} = \frac{\text{number of positive samples}}{\text{number of negative samples}} \tag{2}$$

$$\text{Accuracy (\%)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \times 100 \tag{3}$$

$$\text{Weighted accuracy (\%)} = \frac{\text{TP} + \text{TN} \times \text{IR}}{\text{TP} + \text{FN} + (\text{FP} + \text{TN}) \times \text{IR}} \times 100 \tag{4}$$

$$\text{Precision (\%)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \tag{5}$$

$$\text{Weighted precision (\%)} = \frac{\text{TP}}{\text{TP} + \text{FP} \times \text{IR}} \times 100 \tag{6}$$

$$\begin{aligned}\text{Recall (\%)} &= \text{true positive rate (\%)} \\ &= \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\end{aligned} \tag{7}$$

$$\text{Specificity (\%)} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \tag{8}$$

$$\text{False positive rate (\%)} = \frac{\text{FP}}{\text{TN} + \text{FP}} \times 100 \tag{9}$$

$$\text{F1 score (\%)} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{10}$$

$$\text{Weighted F1 score (\%)} = 2 \times \frac{\text{weighted precision} \times \text{recall}}{\text{weighted precision} + \text{recall}} \tag{11}$$

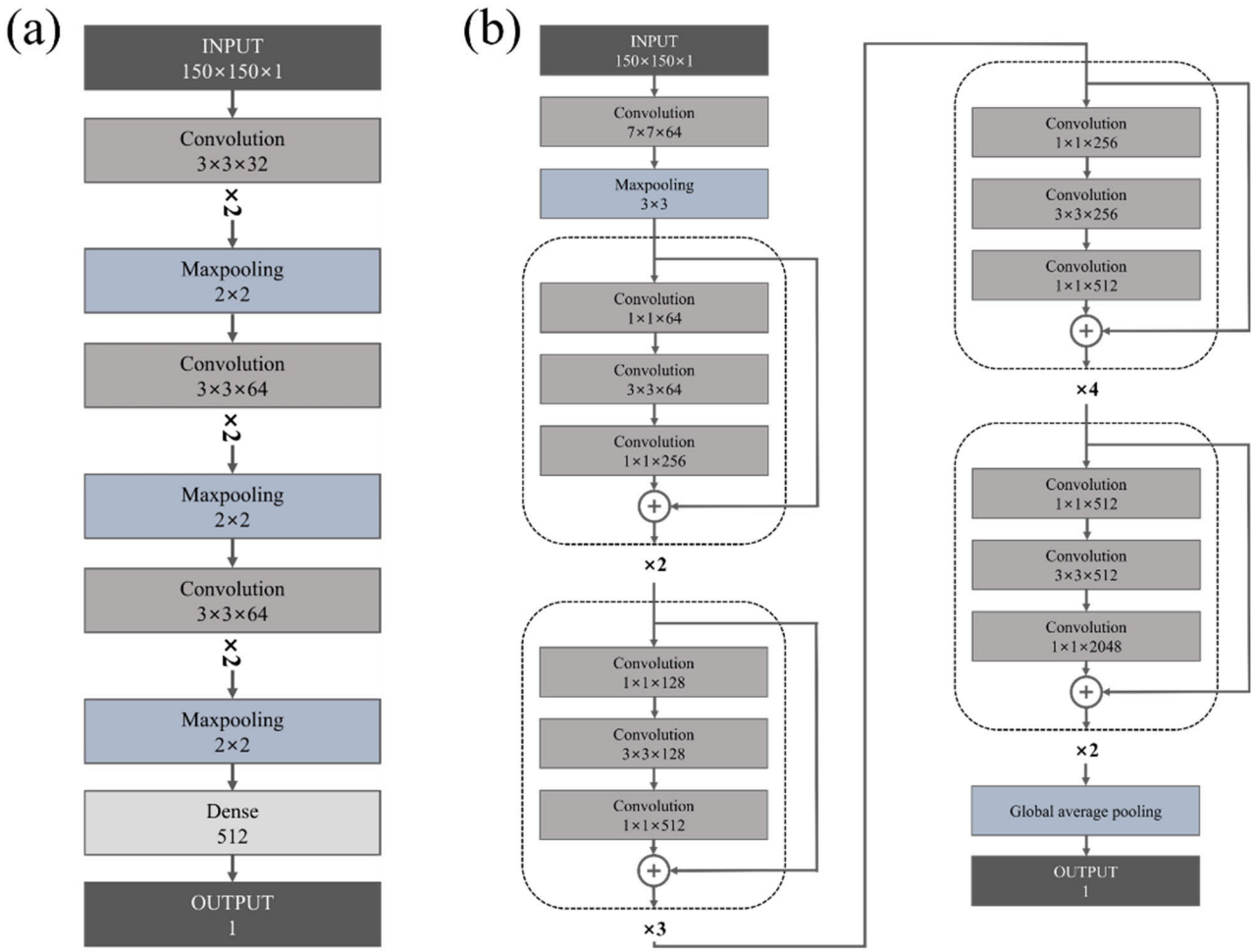

FIGURE 3. Architectures of the (a) VGGNet-based model and (b) ResNet-based model.

TABLE 1. Binary classification criteria.

Criterion	Positive class	Negative class
Cut1	First grade	Second grade Abnormal Non-germinated
Cut2	First grade Second grade	Abnormal Non-germinated
Cut3	First grade Second grade Abnormal	Non-germinated

TABLE 2. Viability test results.

Class	TY sweetiny		Tiniup		Total	
	No.	%	No.	%	No.	%
First-grade	1,021	88.8	272	47.4	1,293	75
Second-grade	36	3.1	149	26.0	185	10.7
Abnormal	42	3.7	103	17.9	145	8.4
Non-germinated	51	4.4	50	8.7	101	5.9
Total	1,150	100.0	574	100.0	1,724	100.0

IV. RESULTS AND DISCUSSION

A. VIABILITY TEST

Table 2 depicts the viability test results for tomato seed samples. Both the TY sweetiny and Tiniup cultivar lots had non-germination rates of <10%, i.e., 4.4% and 8.7%, respectively. The first-grade seeds accounted for 88.8% of the TY sweetiny cultivar, a higher percentage compared to 47.4% for the Tiniup cultivar. The percentages of the second-grade and abnormal classes were 3.1% and 3.7% for the TY sweetiny cultivar and 26% and 17.9% for the Tiniup cultivar, respectively.

B. X-RAY IMAGING

Fig. 4 depicts the X-ray images of tomato seed samples by class. In the non-germinated class, some seeds exhibited

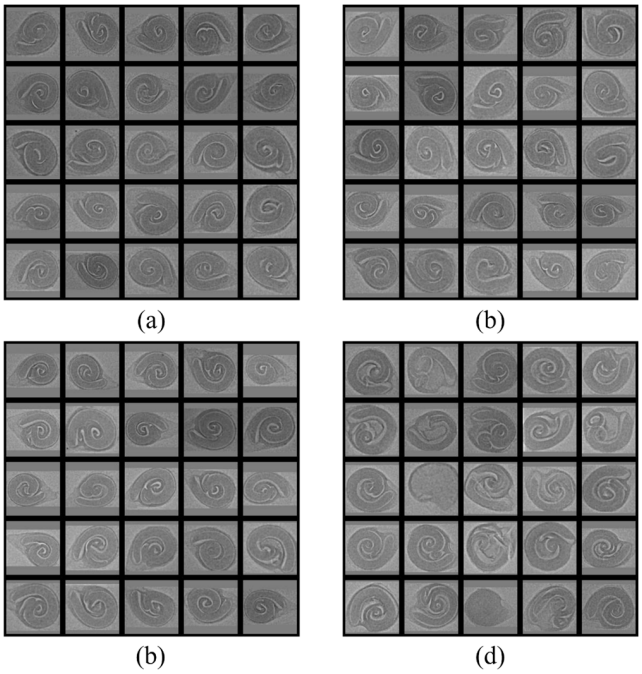

FIGURE 4. X-ray images of tomatoes in four classes: (a) first-grade; (b) second-grade; (c) abnormal; (d) non-germinated.

TABLE 3. Mean and standard deviation of tomato seed integrity (%) quantified via image processing.

Cultivar	Class			
	First-grade	Second - grade	Abnormal	Non-germinated
TY sweetiny	91.61 ± 3.53	87.56 ± 4.12	89.79 ± 3.92	87.33 ± 5.75
Tiniup	87.26 ± 3.26	87.89 ± 3.32	87.56 ± 3.65	81.49 ± 6.72
Total	90.69 ± 3.90	87.84 ± 3.49	88.21 ± 3.86	84.44 ± 6.88

distorted internal structures different from the swirl-shaped internal structure observed in the other classes. As indicated in previous studies [26], [27], [37], the high-intensity area (white area) inside seeds in X-ray images indicates internal cavities. Despite large seed deviations, non-germinated seeds tended to have larger internal cavity areas than in other classes, as indicated in the X-ray images.

C. SEED INTEGRITY

Table 3 and Fig. 5 present the quantified integrity results by cultivar and class. In both cultivars, the non-germinated seeds revealed lower mean integrity than the other seed classes. In the TY sweetiny cultivar, the first-grade class showed a significant difference in seed integrity compared to that in the other classes. The seed integrity of the non-germinated class differed significantly from that of the first-grade and abnormal classes but not from that of the second-grade class. In the Tiniup cultivar, the non-germinated class revealed significant differences from the other classes (Tukey–Kramer test ($p < 0.05$)).

D. INTEGRITY-BASED MODELING

Table 4 presents the evaluation metrics of the integrity-based viability prediction model using the test set. The optimal integrity threshold was determined as the threshold with the highest weighted accuracy for a dataset that included both the training and validation sets. When the integrity of the seed was above the threshold, the seed was classified as positive; otherwise, the seed was classified as negative. The cut3-tested result (i.e., distinguishing germinated and non-germinated seeds) showed the highest evaluation metrics among the test results. Fig. 6 depicts the receiver operating characteristic (ROC) curves of the integrity-based viability prediction model.

E. CNN-BASED MODELING

Table 5 presents the evaluation test results of the CNN models on the test set. Both the VGGNet- and ResNet-based CNN attained higher values in the evaluation metrics for the cut3 model than the cut1 and cut2 models. Most of the cut1 and cut2 models had accuracies ranging from 55.98%–59.77%, while the ResNet-based cut1 model had an accuracy of 67.06%. However, the ResNet-based cut1 model revealed a biased result by cultivar. Therefore, the evaluation results for each cultivar showed accuracies of <50%. In comparison, the VGGNet-based cut3 model had an accuracy of 86.01% and an F1 score of 92.11%; the corresponding values for the ResNet-based cut3 model were 74.33% and 84.42%.

In the weighted metrics, the VGGNet-based cut3 model had an accuracy of 80.92% and an F1 score of 81.14%; the corresponding values for the ResNet-based cut3 model were 76.98% and 76.27%. The cut3-tested ROAUC was similar between the VGGNet-based (84.79%) and ResNet-based (83.31%) models. For these two cut3 models, the unweighted accuracy showed a difference of 11.68%, whereas the difference in the weighted accuracy was 3.94%. Furthermore, the cut3-tested ROAUC difference between the two cut3 models was 1.5%. This can be explained by the greater tendency of the VGGNet-based model to classify seeds as germinated seeds (i.e., the majority class) than that of the Resnet-based model. Therefore, the differences in performance between the models were relatively small considering the class balance, although the VGGNet-based model performed better overall. Fig. 7 depicts the ROC curves of the VGGNet-based models. The cut2- and cut3-tested ROC curves showed similar trends and ROAUC in the cut2 and cut3 models. It is inferred that the abnormal class did not significantly affect training regardless of which binary class it belonged to.

The VGGNet-based model outperformed the integrity-based model by 6.29% in weighted accuracy, 5.39% in weighted F1 score, and 2.66% in cut3 ROAUC. These results indicate that integrity can be a significant factor in predicting viability. However, CNNs that can reflect other features may make more accurate viability predictions than an

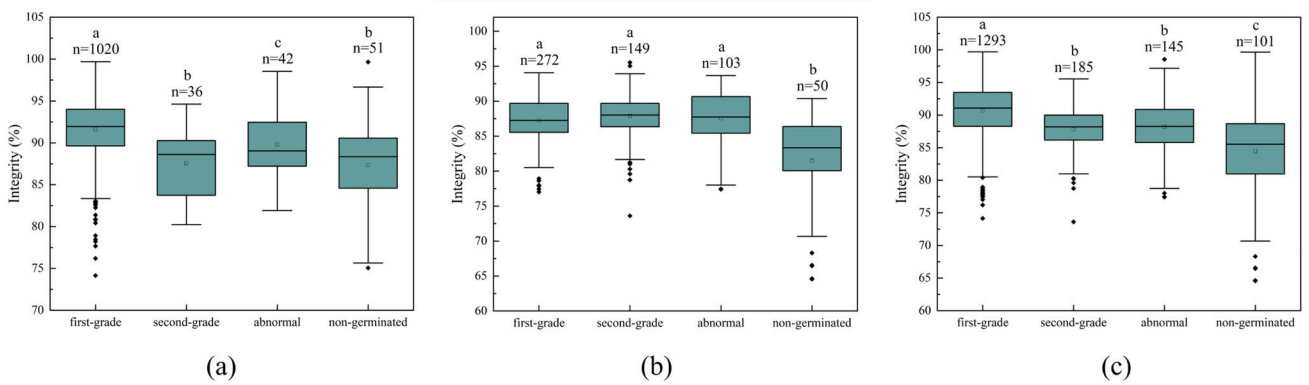

FIGURE 5. Boxplot of seed integrity by viability class: (a) "TY sweetiny" cultivar samples, (b) "Tiniup" cultivar samples, and (c) all samples. Letters indicate statistical differences, as determined using the Tukey–Kramer test ($p < 0.05$).

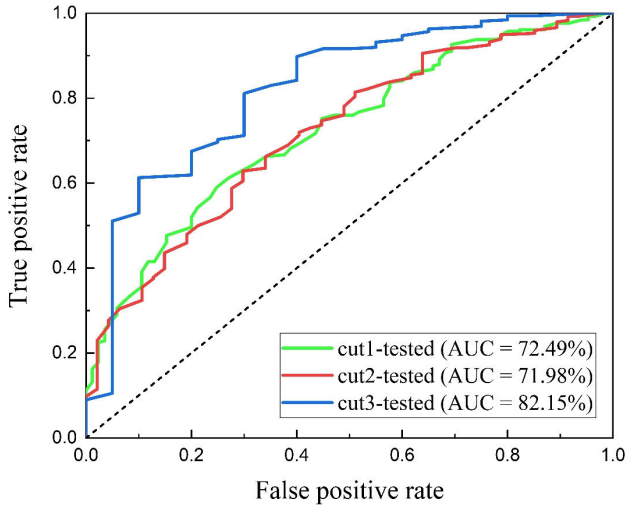

FIGURE 6. Receiver operating characteristic (ROC) curves of the integrity-based model. AUC = area under the curve. Cut1,2,3-tested means that the metric was tested with a binary class test dataset classified following the cut1,2,3 criterion, respectively.

integrity-based model. Considering that the performance difference between the models using integrity alone and the CNN was 2.66% and 6.29%, respectively, a model with a moderate inductive bias could demonstrate similar performance at a lower cost than a CNN. Therefore, cost optimization using various feature extraction and machine learning methods should be attempted in future studies. However, to be generalized for conditions, a CNN capable of learning multiple features suitable for the task is still considered advantageous.

Fig. 8 presents X-ray images of the test set and the VGGNet-based cut3 model prediction results. Images with the highest confidence scores from positive predictions and images with the lowest scores from negative predictions are shown in the figure. The true-positive images with high confidence scores of >97% showed embryo and endosperm structures with small cavity areas. In contrast, the false-negative and true-negative images revealed large cavity areas inside the seeds. Furthermore, the cavity areas of false positive images were smaller than those of negative prediction images. This shows that seed integrity is one of the key features for predicting tomato seed viability in the CNN model. The structures of the embryos and endosperm were not visible in the false-positive image shown in the bottom left of Fig. 8. Notably, among the 1,724 seeds used in our experiment, 12 seeds with no visible internal structures were identified. Therefore, additional data or methods may be required to analyze and model such seeds.

F. EVALUATION BY CULTIVAR

Table 6 presents the cut3-tested results of the integrity- and the VGGNet-based cut3 model for the different seed cultivars. In all evaluation metrics for both cultivars, the VGGNet-based model attained equal or higher values than the integrity-based model. Notably, in both the models, the weighted evaluation metrics and ROAUC values of the Tiniup cultivar were higher than those of the TY sweetiny cultivar, demonstrating that the two models had similar prediction tendencies. For the TY sweetiny cultivar, the VGGNet-based model had an accuracy of 88.21% and an F1 score of 93.56%, while it showed a weighted accuracy of 74.75% and a weighted F1 score of 78%. This result means that the prediction performance for the minority class (non-germinated seeds) is lower than that of the majority class (germinated seeds). In addition, the model has difficulty classifying non-germinated samples of the TY sweetiny cultivar. Considering that the TY sweetiny cultivar has a lower germination rate than the Tiniup cultivar, it is hypothesized that non-germinated samples of the TY sweetiny cultivar deteriorated more than those of the Tiniup cultivar. Consequently, the model could more clearly classify the non-germinated seeds of the Tiniup cultivar.

These results show that conditions of seed lots, such as the deterioration and cultivar, can affect the accuracy and tendency of the model. In a previous viability prediction study for pepper seed [44], the trends differed between seed lots, similar to the results of this study. Since this study

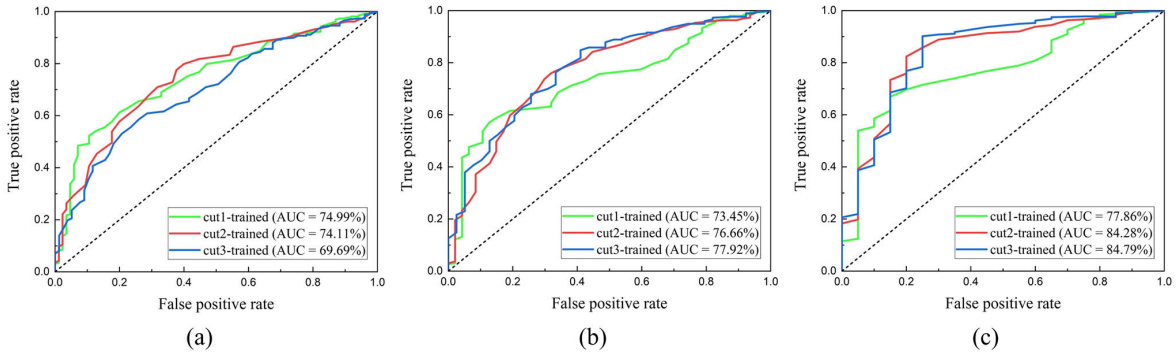

FIGURE 7. Receiver operating characteristic (ROC) curves for VGGNet-based viability prediction models based on the (a) cut1, (b) cut2, and (c) cut3 test sets. AUC = area under the curve. Cut1,2,3-trained means that model was trained with a binary class test dataset classified following the cut1,2,3 criterion.

TABLE 4. Results of the integrity-based model evaluation.

Model	Test criterion	Accuracy (weighted)	F1 score (weighted)	Precision (weighted)	Recall	Specificity	ROAUC
Integrity-based	cut1	72.59% (62.46%)	81.92% (68.74%)	81.30% (58.89%)	82.56%	42.35%	72.49%
	cut2	76.09% (64.67%)	85.30% (69.47%)	90.84% (61.16%)	80.41%	48.94%	71.98%
	cut3	78.72% (74.63%)	87.52% (75.75%)	97.71% (72.54%)	79.26%	70.00%	82.15%

TABLE 5. Results of the convolutional neural network (CNN) model evaluation.

Model		Accuracy (weighted)	F1 score (weighted)	Precision (weighted)	Recall	Specificity	Test criterion	ROAUC
VGGNet-based CNN	cut1	59.77% (69.55%)	65.15% (62.33%)	93.48% (82.75%)	50.00%	89.41%	cut1	74.99%
							cut2	73.45%
							cut3	77.86%
	cut2	58.89% (68.11%)	69.94% (63.50%)	94.80% (74.36%)	55.41%	80.85%	cut1	74.71%
							cut2	76.66%
							cut3	84.28%
	cut3	86.01% (80.92%)	92.11% (81.14%)	98.25% (78.05%)	86.69%	75.00%	cut1	69.44%
							cut2	77.92%
							cut3	84.81%
ResNet-based CNN	cut1	67.06% (70.55%)	74.38% (68.51%)	89.62% (74.28%)	63.57%	77.65%	cut1	74.68%
							cut2	70.26%
							cut3	63.69%
	cut2	55.98% (60.17%)	68.08% (57.76%)	90.96% (61.56%)	54.39%	65.96%	cut1	67.15%
							cut2	67.14%
							cut3	69.58%
	cut3	74.33% (76.98%)	84.42% (76.27%)	98.31% (78.72%)	73.97%	80.00%	cut1	71.47%
							cut2	76.84%
							cut3	83.31%

was limited to two seed lots, the models' results might be biased. Therefore, additional experiments on seed lots under various conditions are required to achieve robust and general models.

TABLE 6. Cut3-tested results by tomato cultivar.

Model	Cultivar	Accuracy (weighted)	F1 score (weighted)	Precision (weighted)	Recall	Specificity	ROAUC
Integrity-based	TY sweetiny	84.28% (67.92%)	91.26% (72.80%)	98.00% (69.11%)	85.84%	50.00%	78.68%
	Tiniup	67.54% (77.69%)	78.61% (74.56%)	98.55% (86.73%)	65.38%	90.00%	85.91%
VGGnet-based CNN	TY sweetiny	88.21% (74.75%)	93.56% (78.00%)	98.00% (69.11%)	89.50%	60.00%	81.23%
	Tiniup	81.58% (85.38%)	88.89% (84.68%)	98.75% (88.96%)	80.61%	90.00%	89.23%

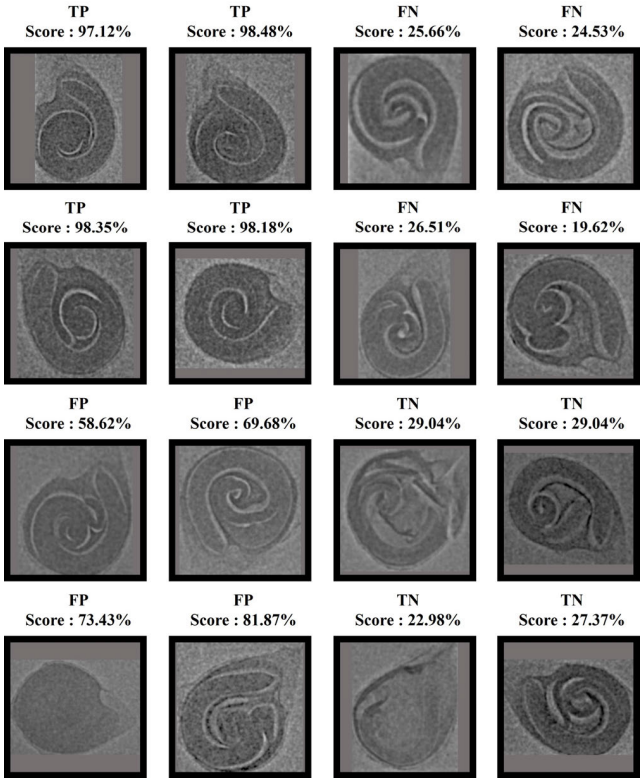

FIGURE 8. Test set X-ray images of tomato seeds and their prediction results obtained using the VGGnet-based cut3 model.

V. CONCLUSION

In this study, viability prediction technologies for tomato seeds were developed using X-ray imaging. The integrity- and CNN-based models were developed and evaluated using multiple viability criteria. The seed integrities were quantified using image processing, and the evaluation results of CNN and integrity-based models were compared. The models performed better in classifying germinated and non-germinated seeds than for the other criteria, and the CNN-based models attained higher evaluation metrics than the integrity-based model. These results show the ability of the CNN-based models to use not only the integrity feature but also other features relevant to germination, including distortion of the seed's internal structure.

Previous studies on viability assessment of tomato seeds through X-ray imaging did not include quantitative prediction modeling. Hyperspectral imaging was mainly used in studies to predict the viability of tomato seeds based on imaging. Shrestha et al. [45] investigated predicting tomato seed viability through hyperspectral imaging; however, no significant differences were detected. Peng et al. [46] achieved accuracies above 85% for tomato seed viability prediction using hyperspectral imaging; however, the accuracy was evaluated for only about 40 seeds. In our study, viable and non-viable seeds could be distinguished with an accuracy of 86.01% using the X-ray imaging method. Furthermore, as the relationship between internal structure and viability of tomato seeds has been reported several times in previous studies, it is considered a robust method. In addition, X-ray imaging has the potential to be applied to seed quality assessment since it can detect internal damage to seeds, which was not included in the current research samples.

Seeds in their natural state that were not artificially aged were used for our experiment and modeling. Therefore, the classes other than the first-grade class were relatively minority classes. In addition, the results of each seed lot confirmed that the model results could differ depending on the condition of the seed lot. Therefore, it is considered that the performance and robustness of the model can be improved through the additional collection of minority class data and data for various seed lots via additional experiments.

REFERENCES

[1] M. M. Rafi, P. N. Yadav, and M. Reyes, "Lycopene inhibits LPS-induced proinflammatory mediator inducible nitric oxide synthase in mouse macrophage cells," *J. Food Sci.*, vol. 72, no. 1, pp. S069–S074, Jan. 2007, doi: 10.1111/J.1750-3841.2006.00219.X.

[2] C. Scolastici, R. O. A. de Lima, L. F. Barbisan, A. L. Ferreira, D. A. Ribeiro, and D. M. F. Salvadori, "Lycopene activity against chemically induced DNA damage in Chinese hamster ovary cells," *Toxicology Vitro*, vol. 21, no. 5, pp. 840–845, Aug. 2007, doi: 10.1016/J.TIV.2007.01.020.

[3] C. Scolastici, R. O. Alves de Lima, L. F. Barbisan, A. L. A. Ferreira, D. A. Ribeiro, and D. M. F. Salvadori, "Antigenotoxicity and antimutagenicity of lycopene in HepG2 cell line evaluated by the comet assay and micronucleus test," *Toxicology Vitro*, vol. 22, no. 2, pp. 510–514, Mar. 2008, doi: 10.1016/J.TIV.2007.11.002.

[4] F. Fornelli, A. Leone, I. Verdesca, F. Minervini, and G. Zacheo, "The influence of lycopene on the proliferation of human breast cell line (MCF-7)," *Toxicology Vitro*, vol. 21, no. 2, pp. 217–223, Mar. 2007, doi: 10.1016/J.TIV.2006.09.024.

[5] P. Chaudhary, A. Sharma, B. Singh, and A. K. Nagpal, "Bioactivities of phytochemicals present in tomato," *J. Food Sci. Technol.*, vol. 55, no. 8, pp. 2833–2849, Aug. 2018, doi: 10.1007/S13197-018-3221-Z.

[6] P. Perkins-Veazie, J. K. Collins, and B. Clevidence, "Watermelons and health," *Acta Horticulturae*, vol. 731, pp. 121–128, Jan. 2007.

[7] A. Rahman and B.-K. Cho, "Assessment of seed quality using non-destructive measurement techniques: A review," *Seed Sci. Res.*, vol. 26, no. 4, pp. 285–305, Dec. 2016, doi: 10.1017/S0960258516000234.

[8] L. M. Kandpal, S. Lohumi, M. S. Kim, J.-S. Kang, and B.-K. Cho, "Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds," *Sens. Actuators B., Chem.*, vol. 229, pp. 534–544, Jun. 2016, doi: 10.1016/j.snb.2016.02.015.

[9] R. K. Singh and H. H. Ram, "Inheritance study of soybean seed storability using an accelerated aging test," *Field Crops Res.*, vol. 13, pp. 89–98, Jan. 1986, doi: 10.1016/0378-4290(86)90013-4.

[10] S. Men, L. Yan, J. Liu, H. Qian, and Q. Luo, "A classification method for seed viability assessment with infrared thermography," *Sensors*, vol. 17, no. 4, p. 845, Apr. 2017, doi: 10.3390/S17040845.

[11] D. Mery, I. Lillo, H. Loebel, V. Riffo, A. Soto, A. Cipriano, and J. M. Aguilera, "Automated fish bone detection using X-ray imaging," *J. Food Eng.*, vol. 105, no. 3, pp. 485–492, Aug. 2011, doi: 10.1016/J.JFOODENG.2011.03.007.

[12] C. Kröger, C. M. Bartle, J. G. West, R. W. Purchas, and C. E. Devine, "Meat tenderness evaluation using dual energy X-ray absorptiometry (DEXA)," *Comput. Electron. Agricult.*, vol. 54, no. 2, pp. 93–100, Dec. 2006, doi: 10.1016/J.COMPAG.2006.09.002.

[13] L. G. Divyanth, V. Chelladurai, M. Loganathan, D. S. Jayas, and P. Soni, "Identification of green gram (vigna radiata) grains infested by callosobruchus maculatus through X-ray imaging and GAN-based image augmentation," *J. Biosystems Eng.*, vol. 47, no. 3, pp. 302–317, Sep. 2022.

[14] P. M. Keagy, B. Parvin, and T. F. Schatzki, "Machine recognition of navel orange worm damage in X-ray images of pistachio nuts," *LWT Food Sci. Technol.*, vol. 29, nos. 1–2, pp. 140–145, 1996, doi: 10.1006/FSTL.1996.0019.

[15] S. K. Mathanker, P. R. Weckler, N. Wang, T. Bowser, and N. O. Maness, "Local adaptive thresholding of pecan X-ray images: Reverse water flow method," *Trans. ASABE*, vol. 53, no. 3, pp. 961–969, 2010, doi: 10.13031/2013.30054.

[16] N. Kotwaliwale, P. R. Weckler, G. H. Brusewitz, G. A. Kranzler, and N. O. Maness, "Non-destructive quality determination of pecans using soft X-rays," *Postharvest Biol. Technol.*, vol. 45, no. 3, pp. 372–380, Sep. 2007, doi: 10.1016/J.POSTHARVBIO.2007.03.008.

[17] S. Kim and T. Schatzki, "Detection of pinholes in almonds through X-ray imaging," *Trans. ASAE*, vol. 44, no. 4, p. 997, 2001, doi: 10.13031/2013.6232.

[18] E. S. Jackson, R. P. Haff, and J. Gomez, "Real-time methods for nondestructive detection of pits in fresh cherries," *Proc. Amer. Soc. Agricult. Biol. Engineers Annu. Int. Meeting*, vol. 5, p. 1, 2009, doi: 10.13031/2013.27094.

[19] R. P. Haff, D. C. Slaughter, Y. Sarig, and A. Kader, "X-ray assessment of translucency in pineapple," *J. Food Process. Preservation*, vol. 30, no. 5, pp. 527–533, Oct. 2006, doi: 10.1111/J.1745-4549.2006.00086.X.

[20] M. A. Shahin, E. W. Tollner, R. D. Gitaitis, D. R. Sumner, and B. W. Maw, "Classification of sweet onions based on internal defects using image processing and neural network techniques," *Trans. ASAE*, vol. 45, no. 5, p. 1613, 2002, doi: 10.13031/2013.11046.

[21] T. F. Schatzki, R. P. Haff, R. Young, I. Can, L.-C. Le, and N. Toyofuku, "Defect detection in apples by means of X-ray imaging," *Trans. ASAE*, vol. 40, no. 5, pp. 1407–1415, 1997, doi: 10.13031/2013.21367.

[22] E. W. Tollner, Y.-C. Hung, B. L. Upchurch, and S. E. Prussia, "Relating X-ray absorption to density and water content in apples," *Trans. ASAE*, vol. 35, no. 6, pp. 1921–1928, 1992, doi: 10.13031/2013.28816.

[23] A. Léonard, S. Blacher, C. Nimmol, and S. Devahastin, "Effect of far-infrared radiation assisted drying on microstructure of banana slices: An illustrative use of X-ray microtomography in microstructural evaluation of a food product," *J. Food Eng.*, vol. 85, no. 1, pp. 154–162, Mar. 2008, doi: 10.1016/J.JFOODENG.2007.07.017.

[24] Y. Liu, W. J. van der Burg, J. W. Aartse, R. A. van Zwol, H. Jalink, and R. J. Bino, "X-ray studies on changes in embryo and endosperm morphology during priming and imbibition of tomato seeds," *Seed Sci. Res.*, vol. 3, no. 3, pp. 171–178, Sep. 1993, doi: 10.1017/S0960258500001756.

[25] W. J. van der Burg, J. W. Aartse, R. A. van Zwol, H. Jalink, and R. J. Bino, "Predicting tomato seedling morphology by X-ray analysis of seeds," *J. Amer. Soc. Horticultural Sci.*, vol. 119, no. 2, pp. 258–263, Mar. 1994.

[26] V. N. Silva, S. M. Cicero, and M. Bennett, "Associations between X-ray visualised internal tomato seed morphology and germination," *Seed Sci. Technol.*, vol. 41, no. 2, pp. 225–234, Aug. 2013, doi: 10.15258/SST.2013.41.2.05.

[27] F. G. Gomes-Junior, J. T. Yagushi, U. L. Belini, S. M. Cicero, and M. Tomazello-Filho, "X-ray densitometry to assess internal seed morphology and quality," *Seed Sci. Technol.*, vol. 40, no. 1, pp. 102–107, Apr. 2012, doi: 10.15258/SST.2012.40.1.11.

[28] T. C. Pearson and D. T. Wicklow, "Detection of corn kernels infected by fungi," *Trans. ASABE*, vol. 49, no. 4, pp. 1235–1245, 2006, doi: 10.13031/2013.21723.

[29] C. Karunakaran, D. S. Jayas, and N. D. G. White, "Detection of internal wheat seed infestation by rhyzopertha dominica using X-ray imaging," *J. Stored Products Res.*, vol. 40, no. 5, pp. 507–516, Jan. 2004, doi: 10.1016/J.JSPR.2003.06.003.

[30] D. S. Narvankar, C. B. Singh, D. S. Jayas, and N. D. G. White, "Assessment of soft X-ray imaging for detection of fungal infection in wheat," *Biosystems Eng.*, vol. 103, no. 1, pp. 49–56, May 2009, doi: 10.1016/J.BIOSYSTEMSENG.2009.01.016.

[31] E. W. Tollner, R. D. Gitaitis, K. W. Seebold, and B. W. Maw, "Experiences with a food product X-ray inspection system for classifying onions," *Appl. Eng. Agricult.*, vol. 21, no. 5, pp. 907–912, 2005, doi: 10.13031/2013.19695.

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[33] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[34] M. R. Ahmed, "Classification of watermelon seeds using morphological patterns of X-ray imaging: A comparison of conventional machine learning and deep learning," *Sensors*, vol. 20, no. 23, pp. 1–15, Nov. 2020, doi: 10.3390/s20236753.

[35] A. D. de Medeiros, R. C. Bernardes, L. J. da Silva, B. A. L. de Freitas, D. C. F. dos Santos Dias, and C. B. da Silva, "Deep learning-based approach using X-ray images for classifying crambe abyssinica seed quality," *Ind. Crops Products*, vol. 164, Jun. 2021, Art. no. 113378, doi: 10.1016/j.indcrop.2021.113378.

[36] F. Corbineau, "Markers of seed quality: From present to future," *Seed Sci. Res.*, vol. 22, no. S1, pp. S61–S68, Feb. 2012, doi: 10.1017/S0960258511000419.

[37] B. Gagliardi and J. Marcos-Filho, "Relationship between germination and bell pepper seed structure assessed by the X-ray test," *Scientia Agricola*, vol. 68, no. 4, pp. 411–416, Aug. 2011.

[38] A. C. C. Lara-Fioreze, C. A. Tomaz, S. L. Fioreze, C. Pilon, and M. D. Zanotto, "Genetic diversity among progenies of crambe abyssinica hochst for seed traits," *Ind. Crops Products*, vol. 50, pp. 771–775, Oct. 2013, doi: 10.1016/j.indcrop.2013.07.039.

[39] G. P. D. Souza, I. A. Borges, C. G. S. Benett, É. F. Leão-Araújo, C. R. D. S. Curvélo, and K. S. S. Benett, "Allelopathic effects of purple nutsedge extract on the physiological quality of cabbage and tomato seeds," *J. Agricult. Sci.*, vol. 11, no. 17, p. 260, Oct. 2019, doi: 10.5539/jas.v11n17p260.

[40] A. D. de Medeiros, L. J. da Silva, J. M. da Silva, D. C. F. dos Santos Dias, and M. D. Pereira, "IJCropSeed: An open-access tool for high-throughput analysis of crop seed radiographs," *Comput. Electron. Agricult.*, vol. 175, Aug. 2020, Art. no. 105555, doi: 10.1016/j.compag.2020.105555.

[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, Dec. 2016, pp. 770–778.
[42] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[43] A. Kolesnikov, "Big transfer (bit): General visual representation learning," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 491–507.
[44] S.-J. Hong, S. Park, A. Lee, S.-Y. Kim, E. Kim, C.-H. Lee, and G. Kim, "Nondestructive prediction of pepper seed viability using single and fusion information of hyperspectral and X-ray images," *Sens. Actuators A, Phys.*, vol. 350, Feb. 2023, Art. no. 114151, doi: 10.1016/j.sna.2022.114151.
[45] S. Shrestha, M. Knapič, U. Žibrat, L. C. Deleuran, and R. Gislum, "Single seed near-infrared hyperspectral imaging in determining tomato (Solanum lycopersicum L.) seed quality in association with multivariate data analysis," *Sens. Actuators B, Chem.*, vol. 237, pp. 1027–1034, Dec. 2016.
[46] Y. Peng, F. Zhao, J. Bai, X. Zheng, W. Wang, and Q. Sun, "Detection and classification of tomato seed vitality based on image processing," *Trans. Chin. Soc. Agricult. Machinery*, vol. 49, no. 2, pp. 327–333, 2018.

SUK-JU HONG received the Ph.D. degree in biosystems engineering from Seoul National University, in 2022. He is currently a Researcher with the National Institute of Agricultural Sciences, Rural Development Administration, South Korea. His research interests include the nondestructive analysis of agricultural products, image processing and deep learning applications in agriculture, and post-harvest engineering. His studies mainly deal with the measurement and modeling of agricultural data using spectroscopy, RGB imaging, X-ray imaging, hyperspectral imaging, thermal imaging, and analysis of properties of agricultural products in the post-harvest process.

SEONGMIN PARK received the Ph.D. degree in biosystems engineering from Seoul National University, South Korea, in 2022. He studied quality evaluation technologies for agricultural products using an electronic nose, RGB imaging, and hyperspectral imaging. He is currently a Researcher with the Research Institute of Agriculture and Life Sciences, Seoul National University. His research interest includes the nondestructive analysis of agricultural products.

CHANG-HYUP LEE received the bachelor's degree in biosystems engineering from Seoul National University, South Korea, in 2020, where he is currently pursuing the Ph.D. degree with the Department of Biosystems Engineering. He studied machine vision systems for agricultural processing machines, hardware for multi-sensing, and deep learning-based imaging technology for diagnosing crop diseases. His research interests include the nondestructive testing of agricultural products in post-harvest processes and measurement technologies in smart farming.

SUNGJAY KIM received the bachelor's degree in agricultural economics and rural development and the master's degree in biosystems engineering from Seoul National University, South Korea, in 2020 and 2023, respectively, where he is currently pursuing the Ph.D. degree with the Department of Biosystems Engineering. He studied UAV-based crop water stress measurement using thermal images and deep-learning-based fruit detection on RGB-D images for the machine vision system of a smart farm harvesting robot. His research interests include nondestructive measurement and deep learning applications in agriculture.

SEUNG-WOO ROH received the bachelor's degree in biosystems engineering from Seoul National University, South Korea, in 2021, where he is currently pursuing the Ph.D. degree with the Department of Biosystems Engineering. He studied deep-learning approaches for disease diagnosis of food crops and hyperspectral imaging-based quality prediction of agricultural products. His research interest includes the nondestructive measurement of agricultural information for quality assessment.

NANDITA IRSAULUL NURHISNA received the bachelor's degree in food science and technology from Hasanuddin University, in 2019, and the master's degree in biosystems engineering from Seoul National University, South Korea, in 2023. She is currently a Researcher with the Research Institute of Agriculture and Life Sciences, Seoul National University. Her research interest includes the quality change of agricultural products in storage. She has studied the quality change of agricultural products such as onions and apples during storage and the development of time-series quality prediction technology using machine learning.

GHISEOK KIM received the Ph.D. degree in agricultural machinery engineering from Chungnam National University, South Korea, in 2007. From 2009 to 2011, he was a Research Scientist with the Composite Vehicle Research Center, Michigan State University, USA. From 2011 to 2015, he was a Research Scientist with the Optical Analysis Equipment Development Team, Korea Basic Science Institute. He is currently an Associate Professor with the Department of Biosystems Engineering, Seoul National University, South Korea. His research interests include analysis of physical and engineering properties of agricultural products, development of postharvest process engineering techniques, and research on nondestructive evaluation of agricultural products including meat and fishery.