



Deep learning models for detection and classification of spongy tissue disorder in mango using X-ray images

Patil Rajvardhan Kiran¹ · G. Avinash² · Mrinmoy Ray² · Sapna Nigam³ · Roaf Ahmad Parray¹ 

Received: 15 March 2024 / Accepted: 14 July 2024 / Published online: 24 July 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Alphonso, a renowned mango variety cultivated in India, is commercially valuable for its delightful taste, vibrant saffron color, texture, and prolonged shelf life, attracts global attention. Spongy Tissue (ST) disorder is prevalent in mangoes, especially in the Alphonso variety, leading to spongy and corky texture in the fruit's pulp. This issue has been identified in as much as 30% of mangoes in a single lot, resulting in the entire lot being rejected during export due to the delayed detection of the disorder. Current mango quality assessment involves destructive cutting, causing fruit damage, leading to significant loss. Limited physical checks don't ensure overall batch quality. By considering above challenges present study demonstrates on use of X-ray imaging technique for detecting internal quality of mango non-destructively along with machine learning algorithms for automatic binary classification (Non spongy vs. Spongy). For evaluation 648 original X-ray images of mangoes were captured and augmented to 3888 images using different augmentation techniques. For automated classification purposes deep learning (DL) models were used. For automated classification, DL models were employed and fine-tuned through grid search hyperparameter optimization within a 5-fold cross-validation framework. Among the DL models, namely CNN, ResNet50, AlexNet, VGG16, and VGG19, they achieved notable overall accuracy rates of 87.27%, 89.18%, 90.21%, 91.75% and 95.82%, respectively, on the test dataset. Significantly, the VGG19 model demonstrated superior performance with highest classification accuracy by TukeyHSD test. DL in X-ray imaging shows promise for detecting internal issues in fruits, enhancing inspection capabilities for diverse purposes, such as identifying diseases or damages.

Highlights

- 'Alphonso' is variety of mango that has suffered from spongy tissue malady.
- Potential of X-ray imaging was tested for identifying internal spongy tissue defects.
- CNN, AlexNet, ResNet, VGG16 and VGG19 models were used for automatic classification.
- Models were optimized by hyperparameter tuning and cross-validation technique.
- Deep transfer learning models achieved >90% mango binary classification accuracy.

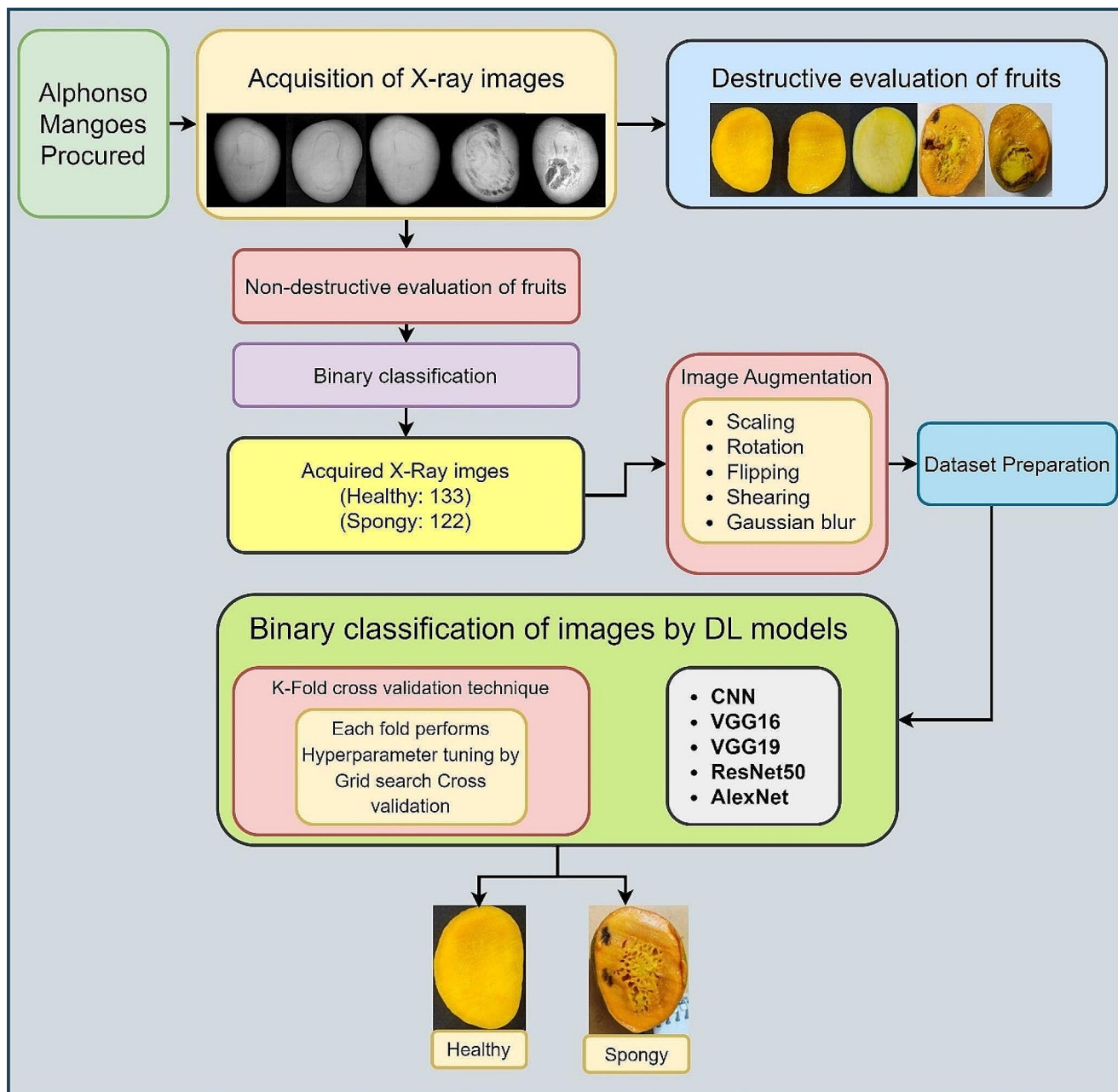
✉ Roaf Ahmad Parray
rouf.engg@gmail.com
Patil Rajvardhan Kiran
rajvardhanp4545@gmail.com
G. Avinash
avinash143stat@gmail.com
Mrinmoy Ray
mrinmoy.ray@icar.gov.in
Sapna Nigam
sapna.nigam@icar.gov.in

¹ Present address: Division of Agricultural Engineering, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India

² Division of Agricultural Statistics, ICAR-Indian Agricultural Statistical Research Institute, New Delhi 110012, India

³ Division of Computer Applications, ICAR-Indian Agricultural Statistical Research Institute, New Delhi 110012, India

Graphical Abstract



Keywords Alphonso mango · Convolutional neural network · Hyperparameter tuning · K-fold cross validation · Radiography

Abbreviations

CNN Convolutional Neural Network
 VGGNet Visual Geometry Group network
 DL Deep Learning
 MAs Milliampere-seconds
 kV Kilovolts
 mA Milliampere

ReLU Rectified Linear Unit
 ResNet Residual Neural Network
 ANN Artificial Neural Network
 STA Spongy Tissue affected
 % Percentage

Introduction

Mango (*Mangifera indica* L) is a highly sought-after fruit cultivated in tropical countries worldwide due to its delectable taste and enticing aroma [1]. India stands as the largest mango producer globally, with the ‘Alphonso’ cultivar being renowned for its distinct flavor and higher cost. However, major mango-growing regions experience internal physiological disorders like Spongy Tissue (ST), jelly seed, insidious fruit rot, tip pulp, and soft nose [2–4]. The ‘Alphonso’ and ‘Keitt’ cultivars, famous for their taste, but suffer annual production losses of 20–30% and 10%, respectively, due to ST disorder [2, 5]. The symptoms usually initiate on the stone’s surface and spread to the mango’s skin, affecting its nutritional value and overall quality [6]. Consequently, stringent phytosanitary regulations in the United States and the European Union have led to a decline in the export potential of these cultivars. India’s Alphonso faced import bans in several countries for an extended period [7].

One of the main challenges in addressing quality disputes related to mangoes is that fruits with infected ST do not exhibit visible symptoms externally, and the flaw becomes apparent only upon cutting [8]. Mango fruits are typically harvested at 75–80% maturity to minimize losses and mechanical injuries during the long supply chain. However, these internal disorders often develop during fruit ripening and can also appear during postharvest storage [9]. Consequently, detecting spongy affected fruits at the farm level becomes exceedingly difficult. Traditionally, the investigation of ST in fruit involved cut and open method, with random sampling from the orchard or fruit lot [6]. As a result of these issues, the marketability and popularity potential of Alphonso and Keitt mangoes are declining steadily [10]. Consumers also express doubts when purchasing these fruits and tend to avoid repurchasing. At present, there are methods that allow the non-destructive sorting and grading of mangoes by evaluating their maturity and physico-chemical characteristics. These techniques include VNIR spectroscopy, hyperspectral imaging, and Nuclear Magnetic Resonance (NMR) [11]. This creates the need for individual fruit prediction to confirm physiological disorders at different grading and sorting stages during the supply chain or prior to shipping for export.

Considering the industries’ demand for a non-invasive technique to evaluate mango internal quality, X-ray imaging emerges as a prominent and sustainable solution [11]. While X-ray imaging is commonly used in the medical field, security purposes, and luggage screening [12], its potential in the food and agricultural industry remains is improving [13]. Soft X-ray imaging offers numerous advantages, including no sample preparation, non-destructiveness, cost-effectiveness, reliability, rapidity, ease of handling, and

high throughput, producing images in just 2 to 3 s [14, 15]. Several previous studies have successfully demonstrated the adaptability of X-ray imaging in the horticulture sector, such as detecting internal defects and bitter rot in apples [16, 17], internal rot in avocados [18], internal disorders in pears [19], seed spoilage location in mangoes [1]. Upon reviewing recent studies, the utilization of X-ray imaging for detecting ST in mangoes has emerged as a significant contribution. Consequently, this technique emerges as a new troubleshooter in the fruit export business, ensuring an uninterrupted supply chain.

In recent years, researchers have increasingly adopted a machine-learning approach to address real-world issues. DL, a subfield of machine learning known for its high capacity to abstract and recognize patterns in images, has gained significant popularity [20]. One of the primary techniques in DL is convolutional neural networks (CNNs), which have become widely used in various domains, particularly for image classification tasks, including fruit classification [1, 21]. Enhanced iterations of convolutional neural networks (CNNs), such as ResNet and VGGNet models, have been employed for classification tasks in the agricultural sector, serving diverse purposes. The performance of these models was investigated to understand their performance in these applications [22–24]. DL, like neural networks, uses layered structures for hierarchical feature learning in tasks such as image or language processing. DL is favored for complex, unstructured data due to its ability to learn intricate patterns. The combination of DL techniques and advanced imaging methods is proving to be highly effective in solving practical challenges across various fields.

First and foremost, this study’s primary objective is to implement a non-destructive and non-invasive X-ray imaging technique for detecting internal ST infection in Alphonso mango fruits. Secondly, the study aims to apply several supervised DL models to the mango X-ray image dataset, enabling accurate classification of the fruit into two categories: Non-spongy and internally damaged (Spongy). Additionally, grid search hyperparameter tuning within each fold of the cross-validation technique has been used to train the DL models effectively. Lastly, all DL models were evaluated through statistical tests to determine their comparative performance superiority.

Materials and methods

Mango fruit sample collection and preparation

A sample set consisting of 324 units of Alphonso mango fruit at 75 to 80% maturity (based on color and specific gravity) was collected from commercial orchards in Ratnagiri (16°

59° 39' N, 73° 37' 09" E) and Sindhudurg (16° 10' 53" N, 73° 44' 86" E) districts of Maharashtra state, India, during two harvesting seasons in 2022 and 2023. After harvesting, the mangoes were immediately kept for pre-cooling. Mangoes with clear skin surfaces, without any damage, wounds, or visible injuries, were carefully selected for image acquisition. Subsequently, numbers were assigned to the mango fruits for identification purposes. During the storage period, the samples were stored at room temperature and non-destructively evaluated at every 5-day interval for up to 25 days of storage.

X-ray image acquisition

Mango samples were randomly chosen from storage, naturally rest position on photostimulable phosphor (PSP) plate, and X-ray images were captured (MARS 40–80 Fixed X-ray (Radio & Fluoro), Allengers Ltd, India) with adjustable voltage, time, and current. Dark current and brightness were corrected for noise reduction. Hermetically sealed air-cooled X-ray source was focused on a table. Voltage (30–45 kV), current (150–600 mAs), and time (4–10 mA) were varied for optimal contrast, resulting in 40 kV, 6 mA, and 600 mAs. X-ray images were processed in Digix ECO software (Allengers, Chandigarh, India), then saved as PNG files from an imaging plate (MD4.0T, AGFA, Hilton, NY).

Discrimination of spongy tissue

In the initial stage of this study, a total of 324 mango units were obtained for the purpose of training and validating models. Subsequently, 648 original X-ray images were taken to enable the non-destructive evaluation of these mangoes. At regular intervals of 5 days (0, 5, 10, 15, 20, and 25 days post-harvest), 54 mangoes were chosen randomly for further examination. This involved capturing X-ray images and halving the mangoes from the equator position near the stone to verify the presence of internal disorders, such as ST in Alphonso mangoes, and to validate the results obtained from X-ray images. In this experiment, all mangoes were divided into two categories only: Non-spongy and Spongy. The division was based on visual inspection by an expert panel. The internal assessment of sliced fruits and the identification through X-ray imaging demonstrates the reliability of this technique.

Data augmentation

A total of 648 original X-ray images were acquired and augmented to diversify the dataset using label-preserving transformations. Data augmentation is a common technique to expand datasets, reduce overfitting, and enhance

generalization. By incorporating additional data from underrepresented classes, data augmentation helped minimize the disparity in class representation [25]. In this study, the initial dataset contained images of both Non-spongy and Spongy mangoes. Five augmentation techniques were applied: horizontal flipping (50% chance), Gaussian blur (sigma 0.5–1), Gaussian noise (sigma 0–0.05), scaling on the x-axis (0.5, 1.0) and y-axis (1.5, 1.0), and horizontal shear (–30 to 30 degrees). The chosen specifications were determined through trial and error, as illustrated in Fig. 1.

Dataset preparation

A dataset comprising 3888 images, including original X-ray images and augmented ones, was created. It contained 1944 Non-spongy and 1944 Spongy images, ensuring balance in dataset in both classes. This dataset was used for model training. The data was split into two parts: 80% (3110 images) for training and 20% (778 images) for validation purpose. Within the 80%, four subsets were created for cross-validation, with each fold containing 388/389 Non-spongy and 388/389 Spongy images. Four folds were used for training DL models, and one for validation iteratively for five rounds. This approach ensured an unbiased evaluation of the model's performance on new data, contributing to the reliability of the study. For testing DL models, a dataset consisting 252 original X-ray images was prepared from 126 Alphonso mangoes randomly collected from the market.

Deep learning models and their architecture for classification

DL is a subfield of artificial intelligence that studies and develops algorithms capable of learning from data and making predictions. The X-ray imaging technique is used as a non-destructive method to discover defects of ST in mangoes, with data gathered in the form of images. Classification and regression analysis are two common machine-learning applications used for detecting defects in horticultural products [26]. Models are created to categorize data using classification algorithms, which forecast discrete outcomes. It focuses on automatically identifying patterns in data using computer algorithms and utilizing such patterns for tasks like categorization [27]. For evaluating horticultural product characteristics, including defect identification, numerous learning algorithms have been applied.

Deep neural network architecture of CNN model

The CNN model comprises essential elements: filter bank kernels, convolution, activation, and pooling layers. It's designed with sequential layers, including two

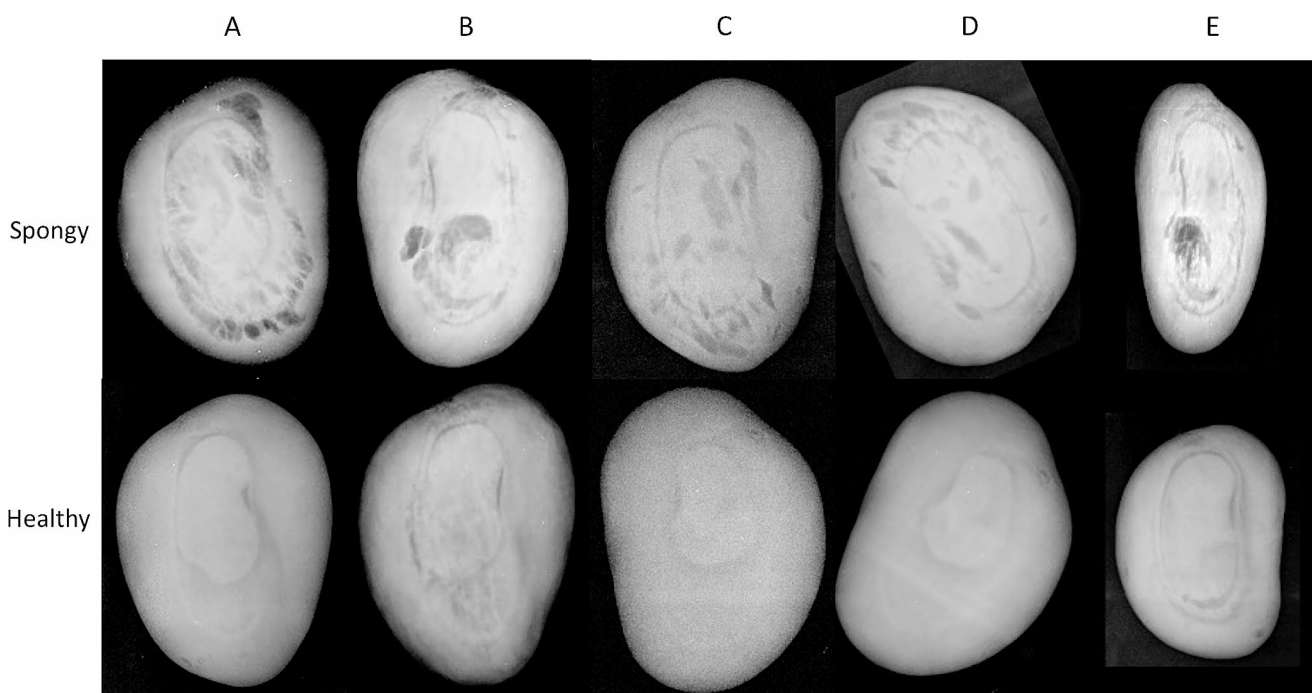


Fig. 1 Augmentation of X-ray images using different techniques (A) Original, (B) Flipping and Gaussian noise, (C) Gaussian blur, (D) Rotation and shearing, (E) Scaling

convolutional layers (32 and 64 filters, 3×3 kernels, ReLU/LeakyReLU activation), followed by max-pooling as shown in Fig. 2. To prevent overfitting, a dropout layer is added. The flattened representation connects to a dense layer (128 neurons, ReLU/LeakyReLU), followed by a final sigmoid dense layer for binary classification. Customizable options include optimizer ('RMSProp' or 'Adam') and learning rate. It's compiled with binary cross-entropy loss and accuracy evaluation, offering adaptability for diverse image classification tasks [21].

Visual geometry group network (VGGNet 16)

In the study, a VGG16 deep neural network was used for mango image classification with 224×224 -pixel input images. The model had five blocks of convolutional layers for feature extraction, with varying numbers of filters in each block. The extracted features were then flattened and connected to a dense layer with two neurons for final classification, outputting probabilities for two classes. This architecture, known for its effectiveness in computer vision tasks, performed well in the study [28].

Visual geometry group network (VGGNet 19)

We tailored a standard VGG19-based neural network for binary classification, keeping VGG19's 16 convolutions, 3 fully connected layers and 3×3 filters. We froze VGG19

base layers, added a 128-neuron dense layer with ReLU activation, and a binary classification layer. Model compiled with binary cross-entropy loss, accuracy metric and optimizer choice.

ResNet

The ResNet model employs skip connections to mitigate the vanishing gradient issue in deep neural networks. It's organized into stages (conv2, conv3, conv4, conv5) with increasing filter counts (64, 128, 256, 512) to capture complex features while reducing spatial dimensions. The final output is flattened to a 100,352-size vector and connected to a Dense layer for binary classification.

AlexNet

The Keras Sequential model follows an AlexNet-style architecture: 96 filters (11×11 , stride 4×4 , ReLU) with Max-Pooling (3×3 , stride 2×2) and Batch Normalization. Then, 256 filters (5×5 , ReLU) with MaxPooling, and three layers with 384 filters (3×3 , ReLU) with MaxPooling and Batch Normalization. Fully connected layers: 4096 units (ReLU, Dropout 0.4, Batch Norm), repeated, and a final layer with 1000 units (ReLU, Dropout 0.4, Batch Norm). Details of all DL models provided in supplementary materials.

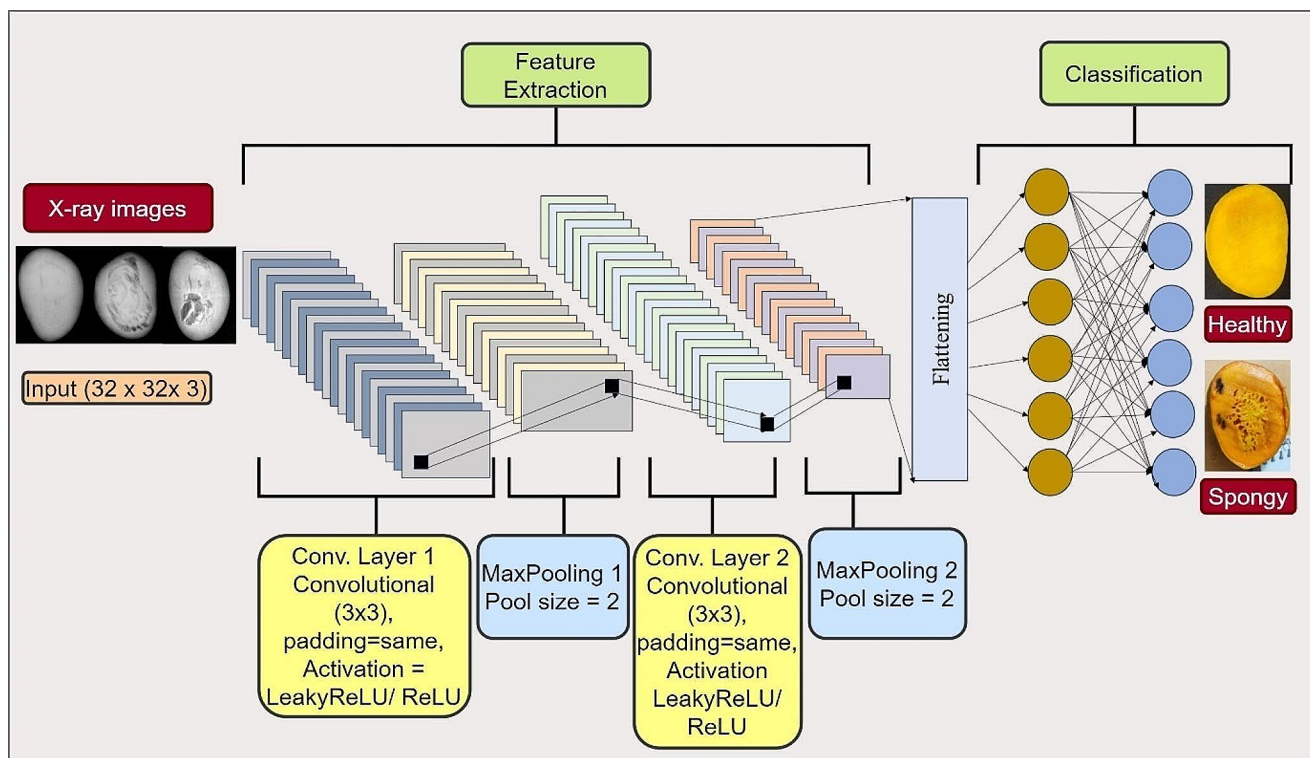


Fig. 2 CNN architecture model used for binary classification of mango

Table 1 Hyperparameters used for training various DL models for image classification

Hyperparameters	Values
Learning rate	0.001, 0.01, 0.1, 0.2
Batch size	32, 64, 128, 256
Dropout rate	0.00, 0.10, 0.20, 0.30
Optimizer	“Adam”, “RMSProp”, “SGD”
Epochs (Stopping Criteria)	30
Activation function	Relu, Tanh, LeakyReLU. Final layer used “Sigmoid” function for classification

Model optimization: Grid search with cross-validation

To address overfitting and underfitting in DL models, we employed grid search hyperparameter tuning. This systematic method optimized various CNN models by adjusting hyperparameters like learning rate, batch size, dropout rate, optimizer type, and early stopping criteria (Table 1). We tested four learning rates (0.001, 0.01, 0.1, and 0.2), four batch sizes (32, 64, 128, 256), four dropout rates (0.00, 0.10, 0.20, and 0.30), and three optimizers (Adam, RMSProp, and SGD). Early stopping with a patience setting of 30 epochs was used to prevent overtraining. Binary cross-entropy loss function and a threshold of 0.95 were applied for binary mango classification. This comprehensive approach resulted in 384 unique hyperparameter combinations per training

fold, ensuring thorough parameter space exploration for optimal model settings.

Following best practices, we employed a 5-fold cross-validation strategy. In each fold, grid search hyperparameter tuning was conducted with 384 combinations on the validation dataset across five folds. This rigorous approach maximized data utilization for training and validation, enhancing model integrity and robustness. It facilitated the exploration of the hyperparameter landscape, ensuring optimized performance and resilience against overfitting [29, 30].

Model selection, validation and testing of model

The performance of each model, corresponding to the varied hyperparameter combinations, was evaluated based on the accuracy across the five folds. This metric provided a dependable estimate of the model’s efficacy with specific hyperparameters, relative to unseen (test) data. The hyperparameter combination that yielded the most favorable average performance across all folds was identified as the optimal model configuration. The assessment of model performance was conducted on a consistent test dataset consisting of 252 images, which remained the same for all models during testing of models.

Model evaluation

The evaluation of model performance on a consistent test dataset comprising 252/253 images, which remained unchanged for all models within the five-fold cross-validation, involved comparing predictions against ground-truth labels of ‘Non-spongy’ (0) and ‘Spongy’ (1). This assessment utilized a confusion matrix to depict true positive (TP), true negative (TN), false positive (FP), and false negative (FN) classifications. From these outcomes, parameters such as accuracy (Eq. 1), precision (Eq. 2), recall (Eq. 3), F1-score (Eq. 4) were computed.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

$$\text{Precision} = TP / (TP + FP) \quad (2)$$

$$\text{Recall} = TP / (TP + FN) \quad (3)$$

$$\text{F1 score} = TP / (TP + 0.5(FP + FN)) \quad (4)$$

Statistical analysis and software

TukeyHSD test assessed mean accuracy differences between models, using a predefined alpha level (usually 0.05) for significance. DL models were developed in Python using TensorFlow, on a Linux system with 32 GB RAM and an NVIDIA GeForce GTX 1080 Ti GPU with 11 GB memory.

Results and discussion

Discrimination of ST and its behavior during storage

X-ray CT imaging stands as an innovative and non-invasive machine vision technique with significant potential for identifying internal flaws in situations where alternative machine vision methods prove inadequate. This technology holds a distinct advantage due to its capacity for X-rays to traverse through the majority of objects, granting it a competitive edge. X-ray images were analyzed to discriminate Non-spongy and ST-affected pulp which exhibits sponginess without apparent external symptoms revealed evident mesocarp damage. This was characterized by distinct dark gray patches on the images, indicating internal air cavities within the flesh. These dark patches were surrounded by light gray areas, representing the regions with healthy pulp. Despite the discernible outline of the endocarp (stone) on the X-ray images, it did not interfere with the identification of ST. In contrast, Non-spongy fruits displayed a consistently light gray color, indicative of uniformly healthy pulp throughout (Fig. 3).

To confirm the presence of internal disorders in X-ray fruit images, samples underwent a destructive method by cutting. On rare occasions, mangoes have been observed to develop spongy tissue (ST) when they reach a maturity level of 75 to 80%, although this phenomenon is not very common at this maturity stage. Detecting ST-affected pulp in early ripening stages is challenging. As the mangoes continue to ripen, the likelihood of spongy tissue occurrence

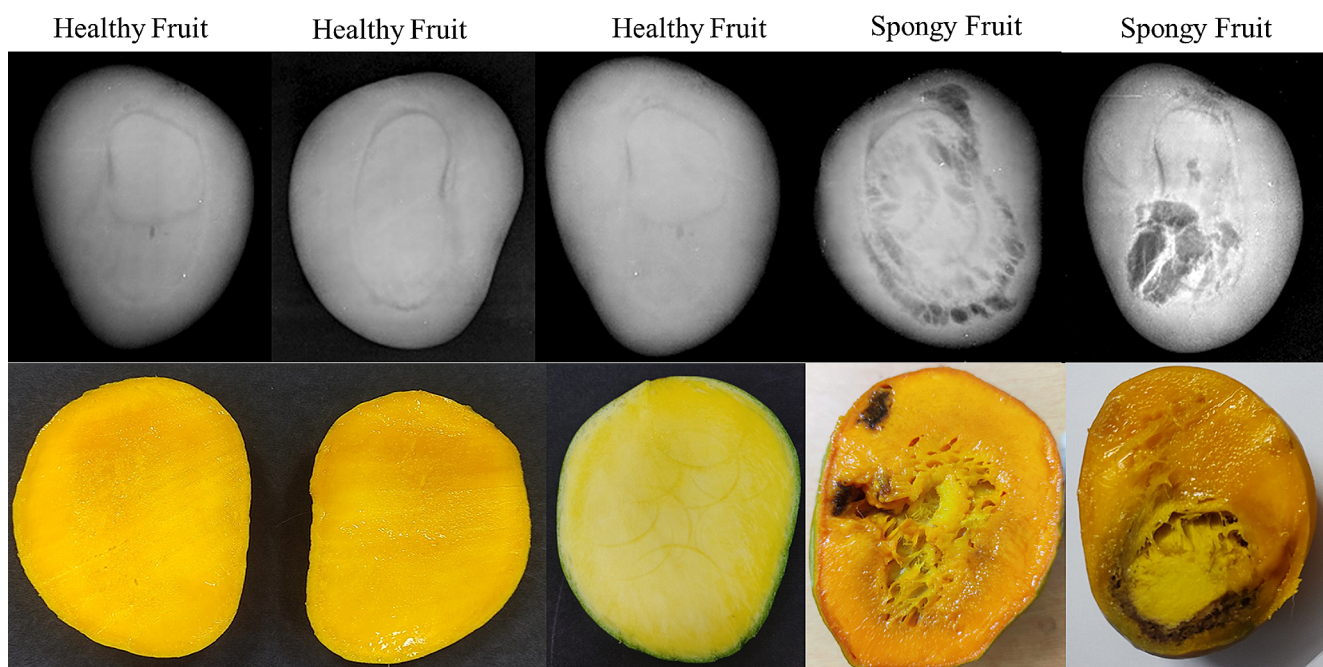


Fig. 3 Discrimination of mango fruits using X-ray imaging and cutting method

increases significantly due to a range of contributing factors. To investigate the progression of spongy tissue development, a research study was conducted where mangoes were subjected to a 25-day storage period, during which alterations in the severity of spongy tissue were observed using X-ray imaging techniques. Subsequent ripening stages witnessed a significant increase in the extent of ST, spreading towards the exocarp of the fruit. The texture and appearance of the ST varied among fruits and underwent changes with the advancing ripening stage. By the conclusion of the storage period, mangoes affected by ST displayed pulp that was highly disintegrated.

In the present study, the assessment of ST focuses solely on its presence, without considering the volume of infection. Notably, there exists morphological variability within the ST. Destructive analyses revealed diverse characteristics of Spongy pulp, including the presence of air pockets with discoloration, discoloration of pulp, white corky and hard pulp, and black soft damaged pulp (Supplementary Fig. S1). Importantly, the nature of the pulp influences the transmittance of X-ray images, introducing variations in the observed characteristics. X-ray imaging was used to determine its effectiveness for internal quality detection of mango fruit. Variations in image grey value result from differential X-ray absorption. The firm fruit pulp exhibits lower grey values due to its increased X-ray absorption, while higher grey values represent reduced X-ray absorption in voids within the fruit and surrounding air. The absorptivity of X-rays in mango pulp is employed to classify it into two categories: Non-spongy and Spongy. Healthy pulp absorbs more X-rays compared to Spongy pulp. So Non-spongy image showed a uniform distribution grey value [11]. Avocado fruits are susceptible to fungal fruit rots during the ripening process. To ensure the quality of the fruits before shipment to retailers, it is crucial to employ efficient and reliable sorting systems [31]. Traditional radiography requires scanning both sides of mangoes due to the interference caused by the presence of the stone center, which can affect X-ray absorbance/transmittance. X-ray CT, offering 3D insights by encompassing entire fruits to reveal internal density variations, stands out as viable method for non-destructive examination

of internal fruit disorders [19, 32]. Despite 3D CT classifiers capturing spatial information, 2D radiography classifiers prove as accurate or even superior (95 to 99% accuracy) compared to 3D models (95 to 96%) [24]. In certain studies, fruits like pears or apples are categorized as consumable or non-consumable [19]. However, when dealing with mangoes affected by ST, such classification is unnecessary. This is due to the aggressive spread of ST, which, once established in the fruit, rapidly absorbs nutrients from the surrounding healthy pulp.

Performance of hyperparameter tuning through grid search

In the hyperparameter tuning of various neural network architectures, distinct preferences emerged across models. The CNN, VGG-19, and ResNet50 architectures demonstrated optimal performance with a learning rate of 0.001, indicating a preference for gradual weight updates. This slower learning rate is often advantageous in complex models or intricate datasets, as it helps in navigating the loss landscape more precisely, avoiding overshooting minima. In contrast, VGG-16 and AlexNet favoured a higher learning rate of 0.01. This higher rate can be indicative of these models' ability to converge faster, possibly due to their less complex structures or different learning dynamics. Interestingly, all models except VGG-16 converged best with a batch size of 128, suggesting a balanced approach to learning that avoids the extremes of very small or large batches. VGG-16's preference for a smaller batch size of 64 could hint at better handling of noise and generalization, as smaller batches often provide more frequent updates and can navigate the loss landscape differently (Table 2).

Augmentation techniques significantly enhance DL performance by expanding limited datasets. These models typically perform better when trained on large datasets, but in certain fields, such extensive data may not be readily available such as X-ray, NMR. Data augmentation addresses this by creating synthetic data through transformations, enlarging datasets, and enhancing model generalization [25, 33]. By dividing X-ray images into training and validation datasets in an 80:20 ratio DL-based algorithms were trained and validated. The division of available data into subsets with similar distributions is essential. In our study dataset was randomly split. Earlier studies employing DL on image data often utilized random splitting [23, 34, 35]. However, random splitting, while appearing fair, may lead to overly optimistic results, as the test set shares the same distribution as the training set. The data split introduces an additional challenge, potentially impacting model performance. To address this, 5-fold cross-validation was implemented in DL, often yielding a model performance that is less biased

Table 2 Results obtained from grid-search hyperparameter tuning with 5-fold cross validation technique for training various DL models

Hyperparameters	CNN	VGG16	VGG19	AlexNet	ResNet50
Learning rate	0.001	0.01	0.001	0.01	0.001
Batch size	128	64	128	128	128
Dropout rate	0.10	0.20	0.20	0.40	0.10
Optimizer	RMSProp	RMSProp	RMSProp	Adam	Adam
Activation function	Leaky-ReLU	ReLU	ReLU	ReLU	ReLU

or optimistic compared to a single random train-validation split. DL automates feature extraction, versatile for produce inspection. Model tuning via grid search optimizes hyperparameters despite complexity [31].

Regarding other parameters, a varied dropout rate was observed: CNN and ResNet50 were tuned to 0.10, indicating a lower regularization need, which might be due to their inherent architectural features. VGG models and AlexNet, with higher dropout rates of 0.20 and 0.40, respectively, suggest a greater need for regularization, possibly to combat overfitting in these architectures. The optimizer choice showed a division: RMSProp was favored for CNN, VGG16, and VGG19, which might be due to its effectiveness in handling the gradients in deeper networks, whereas Adam was chosen for AlexNet and ResNet50, potentially for its adaptive learning rate capabilities. Finally, the activation function was consistently ReLU for all models except the CNN, which opted for LeakyReLU. This choice could reflect the CNN's requirement for addressing the dying ReLU problem, where LeakyReLU helps by allowing a small gradient when the unit is not active.

The industry's key objective is to accurately detect defective fruit to maintain consumer satisfaction and financial viability. Balancing recall and precision are essential for classifier assessment, considering biological variability to

bolster consumer confidence and mango purchases. Evaluating overall model performance, especially for online mango classification, is critical for sustaining supply chains and reducing food waste. Detecting minor defects and preserving healthy fruit minimizes financial losses. Hence, developing models with a high F1 score, such as the VGG19 model in our study, is imperative for mango binary classification, meeting these criteria effectively.

Binary classification of mango using radiography

DL models successfully distinguish Alphonso mango into Non-spongy and spongy classes through X-ray imaging (Fig. 4). CNN model showed a precision of 81.76 and 95.12%, yielding a recall of 95.86 and 79.22%, and showed an F1-score of 0.88 and 0.86 for Non-spongy and spongy classes respectively. Overall accuracy stood at 87.27% on test dataset. AlexNet model achieved a precision of 86.33 and 94.95%, attained recall of 95.39 and 85.07%, and showed a balanced F1-score of 90.91 and 89.71% for Non-spongy and spongy classes respectively. Overall accuracy resulted in 90.21% on test dataset. The ResNet50 model excels in binary image classification, demonstrating 82.5% precision and an impressive 99.19% recall for the "Non-spongy" class, resulting in a balanced F1 score of 90.05%.

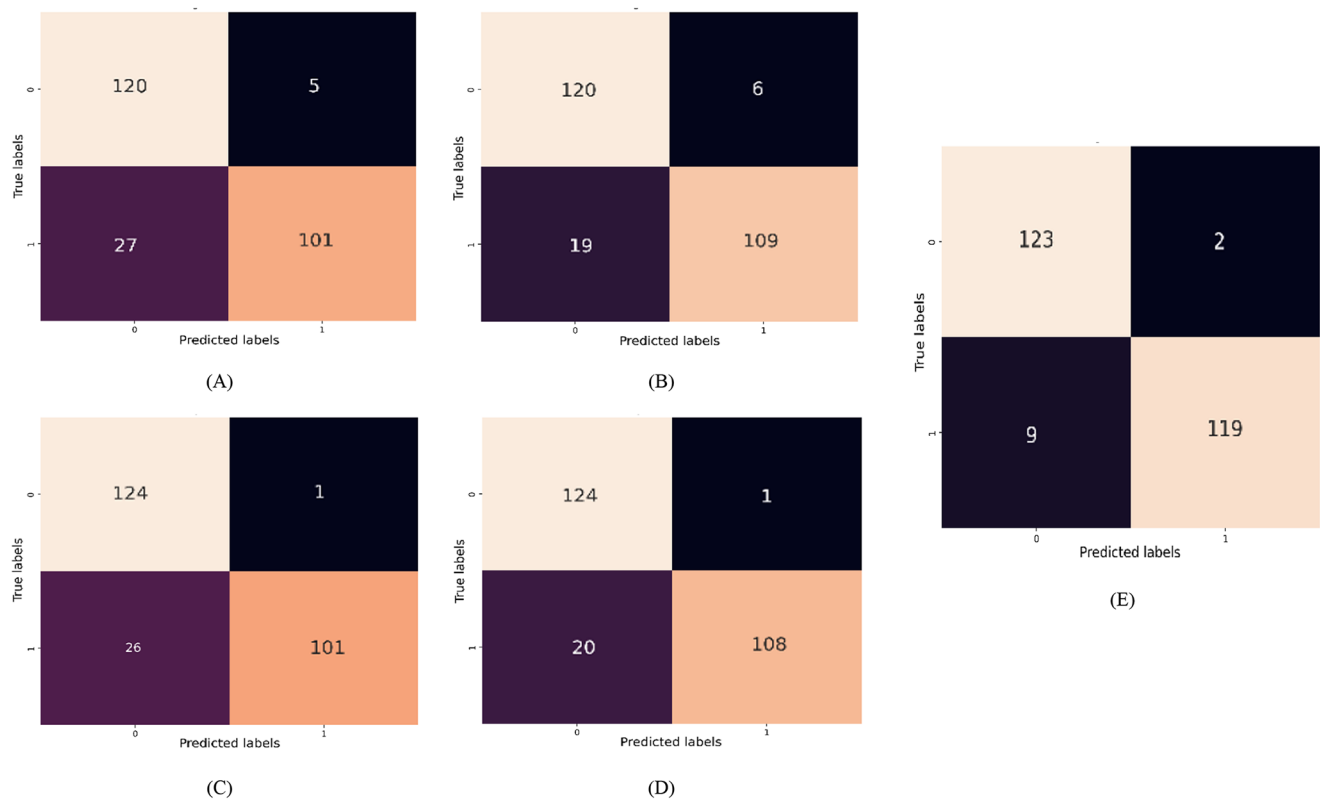


Fig. 4 Performance of various machine learning models on test dataset interpreted by confusion matrix (A) CNN, (B) AlexNet, (C) ResNet50, (D) VGG16, and (E) VGG19

For the “spongy” class, the model achieves exceptional precision at 99.02%, with a strong recall of 79.33%, leading to an effective F1 score of 88.06%. The overall accuracy of the model is reported at 89.18% on test dataset. VGG16 model during binary image classification achieved a precision of 86.95 and 98.76%, recall of 98.91 and 84.91%, and F1 score of 92.25 and 87.58% for Non-spongy and spongy classes respectively. overall accuracy during testing reported 91.75%. VGG19 excels with a 93.47% precision and 98.82% recall for a Non-spongy balanced F1 score of 95.87%. Remarkably, for spongy it achieves a 98.47% precision, 93.14% recall, and an effective F1 score of 95.71%. The overall accuracy stands at 95.82% during testing of model.

After grid search hyperparameter tuning, the best parameters were selected based on validation accuracy. These were used to train the entire dataset for generalization. The optimized hyperparameters were used for testing models (Fig. 5). VGG19 achieved the highest average F1 score for both classes, surpassing other models (VGG16, ResNet50, AlexNet, and CNN). VGG19 also exhibited the lowest standard deviation, indicating consistency. The marginal F1 score difference between “Non-spongy” and “Spongy”

classes suggests comparable accuracy (Fig. 5A). Additionally, VGG19 had the highest recall for both classes, while CNN and ResNet50 showed a significant difference in recall between these classes (Fig. 5B). VGG16 demonstrates the highest precision for “Spongy” fruits, followed by VGG19, whereas for “Non-spongy” fruits, VGG19 outperforms others with a smaller difference between the two classes (Fig. 5C). VGG19 exhibits the highest overall accuracy, followed by VGG16, ResNet50, and AlexNet, while the standard CNN model lags behind (Fig. 5D).

In direct comparison, ResNet50 excelled with the highest recall for the “Non-spongy” class (99.19%) and the highest precision for the “Spongy” class (99.02%), showcasing its strength in accurately identifying “Non-spongy” instances and predicting “Spongy” instances correctly. Meanwhile, VGG19 emerged as the model with the highest overall accuracy (95.82%), underscoring its proficiency in making accurate predictions across both classes. Each model demonstrated unique strengths across precision, recall, F1 score, and overall accuracy, emphasizing the importance of considering these metrics when selecting the most suitable model for a given classification task. All the pairwise combinations from the Tukey HSD test show that there is

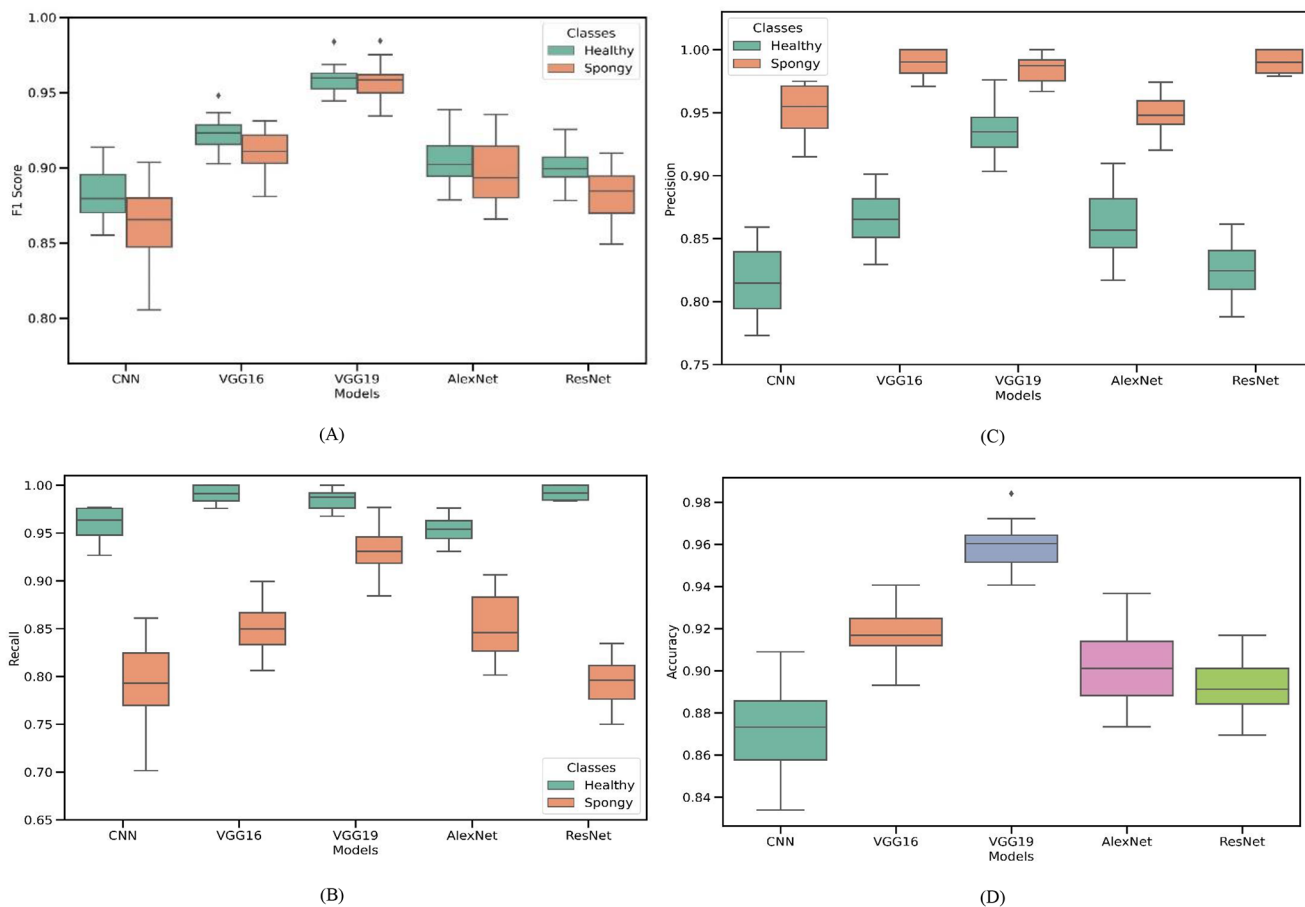
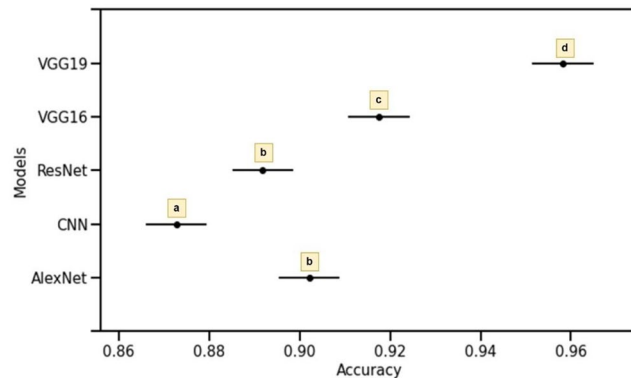


Fig. 5 Comparison of different models using different parameter (A) F1 score, (B) Recall, (C) Precision, (D) Accuracy on test dataset

Table 3 Paired comparison of models

Group1	Group2	meandiff	p-adj
AlexNet	CNN	-0.029	0.001
AlexNet	ResNet50	-0.010	0.222
AlexNet	VGG16	0.015	0.017
AlexNet	VGG19	0.056	0.001
CNN	ResNet50	0.019	0.002
CNN	VGG16	0.044	0.001
CNN	VGG19	0.085	0.001
ResNet50	VGG16	0.025	0.001
ResNet50	VGG19	0.066	0.001
VGG16	VGG19	0.041	0.001

**Fig. 6** Accuracy comparison of model using Tukey HSD test**Table 4** Classification report of various machine learning algorithms on testing dataset

Machine learning model	Fruit class	Precision	Recall	F1 Score	Overall accuracy (%)
CNN	Non-spongy (0)	0.82	0.96	0.88	87.27
	Spongy (1)	0.95	0.79	0.86	
ResNet50	Non-spongy (0)	0.83	0.99	0.90	89.18
	Spongy (1)	0.99	0.79	0.88	
AlexNet	Non-spongy (0)	0.86	0.95	0.91	90.21
	Spongy (1)	0.95	0.85	0.90	
VGG16	Non-spongy (0)	0.87	0.99	0.92	91.75
	Spongy (1)	0.99	0.85	0.88	
VGG19	Non-spongy (0)	0.93	0.99	0.96	95.82
	Spongy (1)	0.98	0.93	0.96	

a significant difference between the models for accuracy parameter except the combination of AlexNet & ResNet50 model. (Table 3; Fig. 6).

In the recent adaptation of the CNN architecture, specifically VGG19, noteworthy classification outcomes were obtained, showcasing exceptional performance. By employing a dataset consisting of 1430 X-ray images of mangoes, the VGG19 model achieved an impressive accuracy rate of 95.82%. Additional metrics like precision, recall, and F1

score also attained a high score of 0.96 for both Non-spongy and Spongy fruits.

In the present study ResNet50 and AlexNet DL models achieved an accuracy of 89.18 and 90.21% on testing dataset. Higher classification accuracy was achieved by both models as compared to CNN model because efficiently learned all the parameters from early activations deeper in the network. For present study precision, recall, F1 score and overall accuracy ratings for all models are given in Table 4. In this study, VGG19 architecture of transfer learning achieved higher accuracy than other architectures. Due to the 5 convolutional layers and increasing filter size 64 to 512, hyperparameters of ReLU activation VGG19 algorithm achieved the best results among all. Similarly, Ansah et al. [1] located seed spoilage in mango fruit using X-ray imaging and deep transfer learning. A VGG 16 DCNN model recorded accuracy of 97.66%. These results highly corroborate with our findings. Van de Looverbosch [19] employed X-ray imaging and deep learning semantic segmentation and utilizing a U-net model for the classification of peach. They achieved 99.4% accuracy for healthy vs. defective peach classification. Matsui et al. [31] used an automated X-ray machine for the detection of internal fruit rot in Hass avocado using DL. They achieved an accuracy score of 0.98 during binary classification using DL-based semantic segmentation.

In our research, a universal threshold value of 0.95 was employed across all models, emphasizing the critical need for accurate classification to uphold export standards. The threshold serves as a decisive marker, categorizing pixels above it as foreground and those below as background. As a result of this approach, our models demonstrated accuracies ranging from 87 to 95%. In contrast, a different study opted for a threshold value of 0.5 [22]. In present study, VGG19 demonstrated superior performance among the models tested and also shows significant result in model accuracies ensured by TukeyHSD test. This highlights the task-specific nature of model performance, leading us to select VGG19 for further classification based on our target and dataset.

Hyperparameter tuning played a significant role in influencing model outcomes. To ensure a fair comparison, model evaluations were conducted on a balanced and sizable dataset using cross-validation techniques. In postharvest research, many studies typically adopt pre-existing model architectures and train them on their unique datasets. A notable challenge in the postharvest domain lies in ensuring the adaptability of models to different varieties, harvesting seasons, storage conditions, and treatments.

The full integration of X-ray imaging in fruit processing industries faces challenges due to the significant initial investment costs and the necessity for a high-voltage power source to generate X-rays. Implementing automated X-ray analysis on existing lines requires significant hardware and

software improvements for quick and continuous scanning. This could streamline internal defect detection on a larger scale, replacing manual analysis. However, challenges include variations in image quality and magnification with different X-ray equipment [1, 24]. Future studies may explore cross-sectional comparisons with diverse equipment and larger datasets to enhance research validity. While the use of this technique raises concerns about the harmful effects of X-rays, effective shielding can mitigate human exposure [12]. Anticipated advancements in instrumentation and computational power, encompassing both hardware and software, suggest that this technique will gain increased applicability in the agriculture field in the future.

Conclusion

This research achieved accurate mango classification for ST presence using DL models. To prevent underfitting/overfitting, thorough hyperparameter tuning was performed individually in each of the five cross-validation folds. This comprehensive approach ensures finely-tuned models with robust performance. The non-destructive detection model introduced here serves as an objective benchmark and can significantly improve post-harvest inspections for Alphonso mangoes. It enhances inspection efficiency, reduces misclassification, and boosts export prospects, fostering consumer trust.

Future research avenues include validating this model's effectiveness with external datasets. This evaluation would assess its resilience when applied to samples with various internal irregularities or mechanical impairments. Such efforts aim to further refine and enhance the system's performance, ensuring its reliability and applicability in diverse mango inspection scenarios.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11694-024-02766-6>.

Acknowledgements We acknowledge the support extended by staff of Indian Agricultural Research Institute (IARI), New Delhi and Indian Council of Agricultural Research, New Delhi for financial support in the form of Senior Research Fellowship.

Author contributions Patil Rajvardhan Kiran: Collecting dataset, Investigation, Analysis, Original draft preparation. G. Avinash: Analysis, Software, Original draft preparation. Mrinmoy Ray: Conceptualization, Supervision. Sapna Nigam: Computer programming. Roaf Ahmad Parray: Conceptualization, Methodology, Review and editing draft.

Funding This research did not receive any specific grant from funding agencies in the public, commercial nor-for-profit sectors.

Data availability The information in this research can be accessed

without any cost. If you wish to utilize the data, please reach out to the corresponding author, Roaf Ahmad Parray, Email: rouf.engg@gmail.com.

Declarations

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. F.A. Ansah, M. Amo-Boateng, E.K. Siabi, P.K. Bordoh, Location of seed spoilage in mango fruit using X-ray imaging and convolutional neural networks. *Sci. Afri.* **20**, e01649 (2023). <https://doi.org/10.1016/j.sciaf.2023.e01649>
2. X. Ma, B. Liu, Y. Zhang, M. Su, B. Zheng, S. Wang, H. Wu, Unravelling correlations between calcium deficiency and ST in mango fruit flesh. *Sci. Hortic.* **309**, 111694 (2023). <https://doi.org/10.1016/j.scienta.2022.111694>
3. S. Shivashankar, V. Ravindra, H.C. Louis, Biochemical changes in seed and mesocarp of mango (*Mangifera indica* L.) cv. 'Alphonso' and their significance during the development of ST. *J. Hortic. Sci. Biotechnol.* **82**(1), 35–40 (2007). <https://doi.org/10.1080/14620316.2007.11512196>
4. L. Raymond, B. Schaffer, J.K. Brecht, J.H. Crane, Internal breakdown in mango fruit: symptomology and histology of jelly seed, soft nose and stem-end cavity. *Postharvest Biol. Technol.* **13**(1), 59–70 (1998). [https://doi.org/10.1016/S0925-5214\(97\)00074-4](https://doi.org/10.1016/S0925-5214(97)00074-4)
5. P. Oak, A. Deshpande, A. Giri, V. Gupta, Metabolomic dynamics reveals oxidative stress in ST disorder during ripening of *Mangifera indica* L. fruit. *Metabolites.* **9**(11), 255 (2019). <https://doi.org/10.3390/metabo9110255>
6. P. Oak, V. Jha, A. Deshpande, R. Tanpure, V. Dawkar, S. Mundhe et al., Transcriptional and translational perturbation in abiotic stress induced physiological activities and metabolic pathway networks in ST disorder of mango fruit. *Postharvest Biol. Technol.* **188**, 111880 (2023). <https://doi.org/10.1016/j.postharvbio.2022.111880>
7. A. Raghavendra, D.S. Guru, M.K. Rao, Mango internal defect detection based on optimal wavelength selection method using NIR spectroscopy. *Artif. Intell. Agric.* **5**, 43–51 (2021). <https://doi.org/10.1016/j.aiaa.2021.01.005>
8. M.T. Janave, A. Sharma, Spongy tissue development in Alphonso mango: association with *Staphylococcus xylosum*. *Eur. J. Plant. Pathol.* **122**, 335–348 (2008). <https://doi.org/10.1007/s10658-008-9298-x>
9. S. Shivashankar, Physiological disorders of mango fruit. *Hortic. Rev.* **42**, 313–348 (2014). <https://doi.org/10.1002/9781118916827.ch06>
10. X. Ma, J. Wang, M. Su, B. Liu, B. Du, Y. Zhang et al., The link between mineral elements variation and internal flesh breakdown of 'Keitt' mango in a steep slope mountain area, Southwest China. *Hortic.* **8**(6), 533 (2022). <https://doi.org/10.3390/horticulturae8060533>
11. N.K. Mahanti, R. Pandiselvam, A. Kothakota, S.K. Chakraborty, M. Kumar, D. Cozzolino, Emerging non-destructive imaging techniques for fruit damage detection: image processing and analysis. *Trends Food Sci. Technol.* **120**, 418–438 (2022). <https://doi.org/10.1016/j.tifs.2021.12.021>
12. N. Kotwaliwale, K. Singh, A. Kalne, S.N. Jha, N. Seth, A. Kar, X-ray imaging methods for internal quality evaluation of

- agricultural produce. *J. Food Sci. Technol.* **51**, 1–15 (2014). <https://doi.org/10.1007/s13197-011-0485-y>
13. M. Azadbakht, M. VahediTorshizi, The use of CT scan imaging technique to determine pear bruise level due to external loads. *Food Sci. Nutr.* **7**(1), 273–280 (2019). <https://doi.org/10.1002/fsn3.882>
 14. Z. Du, Y. Hu, N. Ali Buttar, A. Mahmood, X-Ray computed tomography for quality inspection of agricultural products: a review. *Food Sci. Nutr.* **7**(10), 3146–3160 (2019). <https://doi.org/10.1002/fsn3.1179>
 15. X. Feng, H. Zhang, P. Yu, X-ray fluorescence application in food, feed, and agricultural science: a critical review. *Crit. Rev. Food Sci. Nutr.* 1–11 (2020). <https://doi.org/10.1080/10408398.2020.1776677>
 16. M. Van Dael, P. Verboven, A. Zanella, J. Sijbers, B. Nicolai, Combination of shape and X-ray inspection for apple internal quality control: in silico analysis of the methodology based on X-ray computed tomography. *Postharvest Biol. Technol.* **148**, 218–227 (2019). <https://doi.org/10.1016/j.postharvbio.2018.05.020>
 17. Y. Si, S. Sankaran, Computed tomography imaging-based bitter pit evaluation in apples. *Biosyst Eng.* **151**, 9–16 (2016). <https://doi.org/10.1016/j.biosystemseng.2016.08.008>
 18. T. Matsui, T. Kamata, S. Koseki, K. Koyama, Development of automatic detection model for stem-end rots of ‘Hass’ avocado fruit using X-ray imaging and image processing. *Postharvest Biol. Technol.* **192**, 111996 (2022). <https://doi.org/10.1016/j.postharvbio.2022.111996>
 19. T. Van de Looverbosch, E. Raeymaekers, P. Verboven, J. Sijbers, B. Nicolai, Non-destructive internal disorder detection of Conference pears by semantic segmentation of X-ray CT scans using deep learning. *Expert Sys. Appl.* **176**, 114925 (2021). <https://doi.org/10.1016/j.eswa.2021.114925>
 20. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature.* **521**(7553), 436–444 (2015). <https://doi.org/10.1038/nature14539>
 21. J. Naranjo-Torres, M. Mora, R. Hernández-García, R.J. Barrientos, C. Fredes, A. Valenzuela, A review of convolutional neural network applied to fruit image processing. *Appl. Sci.* **10**(10), 3443 (2020). <https://doi.org/10.3390/app10103443>
 22. A.D.D. Medeiros, L.J.D. Silva, J.P.O. Ribeiro, K.C. Ferreira, J.T.F. Rosas, A.A. Santos, C.B.D. Silva, Machine learning for seed quality classification: an advanced approach using merger data from FT-NIR spectroscopy and X-ray imaging. *Sensors.* **20**(15), 4319 (2020). <https://doi.org/10.3390/s20154319>
 23. A. Fazari, O.J. Pellicer-Valero, J. Gómez-Sanchis, B. Bernardi, S. Cubero, S. Benalia et al., Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images. *Comput. Electron. Agric.* **187**, 106252 (2021). <https://doi.org/10.1016/j.compag.2021.106252>
 24. A. Tempelaere, L. Van Doorselaer, J. He, P. Verboven, B.M. Nicolai, BraeNet: Internal Disorder detection in ‘Braeburn’ apple using X-ray imaging data. *Food Control.* **155**, 110092 (2024). <https://doi.org/10.1016/j.foodcont.2023.110092>
 25. P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, A. Haworth, *J. Med. Imaging Radiat. Oncol.* **65**(5), 545–563 (2021). <https://doi.org/10.1111/1754-9485.13261>. A review of medical image data augmentation techniques for deep learning applications
 26. S. Naik, B. Patel, Machine vision-based fruit classification and grading—a review. *Int. J. Comput. Appl.* **170**(9), 22–34 (2017)
 27. J.F.I. Nturambirwe, U.L. Opara, Machine learning applications to non-destructive defect detection in horticultural products. *Biosys Eng.* **189**, 60–83 (2020). <https://doi.org/10.1016/j.biosystemseng.2019.11.011>
 28. S.A. Althubiti, F. Alenezi, S.K.S. Shitharth, C.V.S. Reddy, Circuit manufacturing defect detection using VGG16 convolutional neural networks. *Wirel. Commun. Mob. Com.* **2022**, 1–10 (2022). <https://doi.org/10.1155/2022/1070405>
 29. H. Chen, Z. Zhang, W. Yin, C. Zhao, F. Wang, Y. Li, A study on depth classification of defects by machine learning based on hyper-parameter search. *Measurement.* **189**, 110660 (2022). <https://doi.org/10.1016/j.measurement.2021.110660>
 30. X. Jiang, C. Xu, Deep learning and machine learning with grid search to predict later occurrence of breast Cancer metastasis using clinical data. *J. Clin. Med.* **11**(19), 5772 (2022). <https://doi.org/10.3390/jcm11195772>
 31. T. Matsui, H. Sugimori, S. Koseki, K. Koyama, Automated detection of internal fruit rot in Hass avocado via deep learning-based semantic segmentation of X-ray images. *Postharvest Biol. Technol.* **203**, 112390 (2023). <https://doi.org/10.1016/j.postharvbio.2023.112390>
 32. E. Arendse, O.A. Fawole, L.S. Magwaza, U.L. Opara, Non-destructive prediction of internal and external quality attributes of fruit with thick rind: a review. *J. Food Eng.* **217**, 11–23 (2018). <https://doi.org/10.1016/j.jfoodeng.2017.08.009>
 33. A. Saber, M. Sakr, O.M. Abo-Seida, A. Keshk, H. Chen, A novel deep-learning model for automatic detection and classification of breast cancer using the transfer learning technique. *IEEE Access.* **9**, 71194–71712 (2021)
 34. X. Liang, X. Jia, W. Huang, X. He, L. Li, S.S. Fan et al., Real-time grading of defect apples using semantic segmentation combination with a pruned YOLO V4 network. *Foods.* **11**(19), 3150 (2022). <https://doi.org/10.3390/foods11193150>
 35. A. Tempelaere, T. Van De Looverbosch, K. Kelchtermans, P. Verboven, T. Tuytelaars, B. Nicolai, Synthetic data for X-ray CT of healthy and disordered pear fruit using deep learning. *Postharvest Biol. Technol.* **200**, 112342 (2023). <https://doi.org/10.1016/j.postharvbio.2023.112342>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.