Smart Agricultural Technology 7 (2024) 100433

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Image processing techniques to identify tomato quality under market conditions

Thilina Abekoon [a], Hirushan Sajindra [a], J . A . D . C . A . Jayakody [b], E . R . J Samarakoon [c], Upaka Rathnayake [d,*]

[a] Water Resources Management and Soft Computing Research Laboratory, Millennium City, Athurugiriya 10150, Sri Lanka
[b] Faculty of Computing, Sri Lanka Institute of Information Technology (SLIIT), New Kandy Road, Malabe 10115, Sri Lanka
[c] Department of Food Science and Technology, Faculty of Agriculture, University of Peradeniya, 20400, Sri Lanka
[d] Department of Civil Engineering and Construction, Faculty of Engineering and Design, Atlantic Technological University, Sligo F91 YW50, Ireland

ARTICLE INFO

Keywords:
Image processing
Post harvest technology
Tomato
Machine learning
Convolutional neural network
Classification

ABSTRACT

Tomatoes are essential in both agriculture and culinary spheres, demanding rigorous quality assessment. It is highly advantageous to discern the maturity level and the time range post-harvesting of tomatoes in the market through visual analysis of their images. This research endeavors to forecast tomato quality by accurately determining the maturity level and the duration post-harvest, specifically tailored to Sri Lankan market conditions, with a particular focus on Padma tomatoes. It identifies maturity stages (Green, Breakers, Turning, Pink, Light Red, Red) and post-harvest dates using image processing techniques. Greenhouse-grown Padma tomatoes mimic market conditions for image capture, and Convolutional Neural Networks facilitate this analysis. Model 1, using ReLU and sigmoid activation functions, accurately classifies tomatoes with 99 % training and validation accuracy. Model 2, with seven classes, achieves 99 % training and 98 % validation accuracy using ReLU and softmax activation functions. Integration of the IPGRI/IITA 1998 classification method enhances tomato categorization. Efficient tomato image screening optimizes resource use. This study successfully determines Padma tomato post-harvest dates based on maturity stages, a significant contribution to tomato quality assessment under market conditions.

Introduction

Tomato (*Solanum lycopersicum* L.) is a major vegetable and herbaceous perennial fruit crop that has gained immense popularity over the last century. It is cultivated in almost every country worldwide in open fields, greenhouses, and net houses [1,2]. Botanically, tomatoes are classified as fruits based on their biological characteristics. They originate from the fertilized ovary of a flower and contain seeds, aligning with the botanical definition of a fruit, which encompasses structures developed from the ovary of a flowering plant. Conversely, from a culinary standpoint, tomatoes are frequently regarded as vegetables. This culinary classification is primarily due to their common usage in savory dishes and their flavor profile, which is more reminiscent of vegetables [3]. Tomatoes contain numerous health-promoting compounds and can be easily integrated as a nutritious part of a balanced diet [4]. It contains various health-promoting compounds, including vitamins, carotenoids, and phenolic compounds, which provide physiological advantages, including anti-inflammatory, anti-allergenic, antimicrobial, vasodilatory, antithrombotic, cardio-protective, and antioxidant effects [5,6]. In addition to being consumed as fresh fruits, tomatoes are also utilized in various processed products such as canned tomatoes, paste, puree, ketchup, juice and pasta sauces [7]. In 2021, the world's leading tomato producers were China, India, Turkey, and the United States of America, with respective productions of 67.67, 21.18, 13.09, and 10.47 Mt [8]. According to the United States Department of Agriculture, tomatoes are classified into four grades: U.S. No. 1, U.S. Combination, U.S. No. 2, and U.S. No. 3. The classification is based on several factors, including maturity, color, shape, size, and cleanliness of the tomatoes. Regarding color classification, tomatoes are divided into several stages: Green, Breakers, Turning, Pink, Light Red, and Red [9] as shown in Fig 1.

During the green stage, the tomato surface is characterized by a

Abbreviations: CNN, Convolutional Neural Network; ReLU, Rectified Linear Unit.

* Corresponding author.
E-mail address: upaka.rathnayake@atu.ie (U. Rathnayake).

https://doi.org/10.1016/j.atech.2024.100433
Received 15 January 2024; Received in revised form 3 March 2024; Accepted 5 March 2024
Available online 6 March 2024

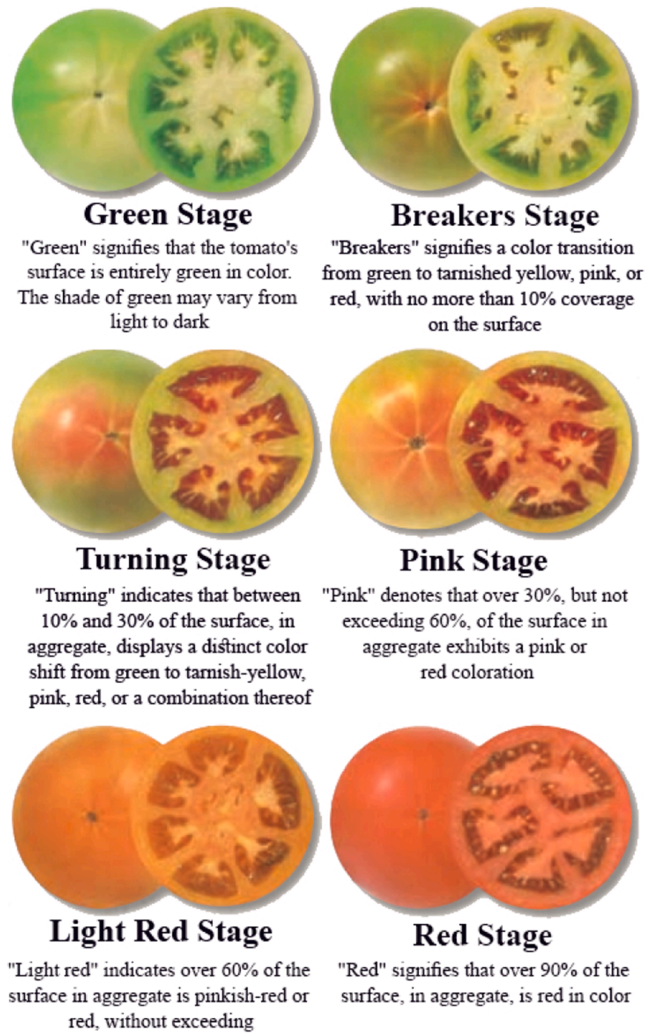

Fig. 1. Maturity and ripening stages of tomatoes based on United States standards for grades of fresh tomatoes (Source: [9]).

vibrant green hue. As it progresses to the breaker stage, subtle hints of yellow, pink, and red emerge, encompassing no more than 10 % of the surface. Transitioning further to the turning stage, the tomato takes on a tarnished-yellow appearance, accompanied by a gradual increase in pink and red, which now extends to cover 10 % to 30 % of the surface. As the tomato enters the pink stage, these color changes become more pronounced, and aggregates begin to exhibit a distinct pink or red coloration, spanning a range of 30 % to 60 % of the surface. Advancing to the light red stage, the tomato's surface transforms into a delicate pinkish-red or rich red hue, now covering more than 60 % of the tomato's surface area. Finally, in the red stage, the tomato's surface becomes predominantly adorned with red color aggregates, extending to encompass more than 90 % of its surface area [9].

In the context of Sri Lanka's Yala and Maha seasons for the year 2022/2023, a noteworthy production of 139,978 t of tomatoes was achieved [10]. There are numerous tomato varieties in the world, and the biochemical compounds and physical characteristics of the tomato vary depending on the specific variety [11]. Within the agricultural landscape of Sri Lanka, the Department of Agriculture has thoughtfully recommended a selection of tomato varieties, each possessing distinct varietal attributes. This assortment includes KWR, T146, T245, Thilina, Ravi, Tharidu, Rashmi, Rajitha, Lanka Sour (*Goraka Takkali*), Maheshi, Bathiya, and Lanka cherry [12].

In Sri Lanka, the major districts for tomato cultivation are Badulla, Nuwara-Eliya, Kandy, and Matale [13]. These districts have garnered significance due to their favorable climatic conditions, fertile soil, and suitable terrain, all of which contribute to the successful cultivation of this vibrant and nutritious crop. The optimal climatic conditions for successful tomato cultivation encompass a temperature range of 21 °C to 24 °C, a soil pH between 5.5 and 7.5, and an elevation of 1000 to 2000 m [14].

Padma is a highly bacterial wilt-resistant tomato variety in tropical region [15,16]. It has recently gained popularity among Sri Lankan farmers and consumers due to its superior keeping quality. The focus of this research is to delve into the intricacies of the Padma tomato variety's quality attributes through the utilization of advanced image processing techniques.

In the tomato industry, tomato quality evaluation is widely employed in processing lines. To be effective, efficient, accurate, and sustainable without human intervention, artificial intelligence techniques are extensively utilized [17,18]. Artificial intelligence driven image processing and computer vision algorithms analyze digital images of tomatoes, identifying crucial attributes such as maturity, size, color, shape, and defects [17,19]. Additionally, machine learning models, including Support Vector Machines and Deep Learning networks like Convolutional Neural Networks, classify tomatoes based on annotated data, enabling precise quality assessment [20,21]. Spectroscopic and hyperspectral imaging techniques offer valuable chemical composition insights [22,23], while internet of things sensors and robotics contribute real-time data and automation capabilities [24]. The fusion of data from diverse sources allows for a comprehensive analysis of tomato quality [25]. As a result, the tomato industry benefits from streamlined sorting and grading processes, ensuring that consumers receive high-quality products with optimal nutritional content.

The utilization of image processing not only evaluates the quality of the tomatoes but also offers an assessment of their shelf life. This comprehensive research involves an extensive exploration of Padma tomato quality, maturity stages, and post-harvest dates based on surface conditions through the use of image processing techniques under market conditions in Sri Lanka. This involved capturing a total of 12342 daily images of Padma tomatoes until they reached a state of deterioration. By providing detailed insights into the RGB color values, the shapes, and the shelf life of Padma tomatoes, this research empowers various stakeholders within the agricultural and culinary sectors. For farmers, it offers valuable information that can help optimize cultivation practices, allowing for better crop management and resource allocation. Processors can benefit from this knowledge by improving their harvesting schedules and processing techniques, ensuring that the tomatoes they handle meet the highest quality standards. Moreover, consumers gain access to information that enables them to make more informed choices about the tomatoes they purchase, ensuring they receive fresh and nutritious produce. Ultimately, this study elevates the broader landscape of tomato quality management, solidifying the Padma tomato variety's position as a preferred choice in both agricultural and culinary contexts.

Materials and methods

Preliminary arrangements for tomato nursery

Initially, Padma tomatoes growing in greenhouses were selected for the study from various locations across Sri Lanka. Greenhouses offer the most advantageous option for achieving both quality and quantity in tomato production. Beyond delivering elevated yields, greenhouse cultivation also provides a sanctuary free from dust, insects, diseases, and pests. Furthermore, the favorable environment within greenhouses ensures uniform fruit size, contributing to a consistent and superior harvest [26]. Subsequently, four distinct locations were chosen from the major tomato cultivation districts (Fig 2) [13]. The selected locations were Walawela in Matale District, Marassana in Kandy District, Rikillagaskada in Nuwara Eliya District and Keppetipola in Badulla District. Twenty-five tomatoes were harvested at the green stage from each of

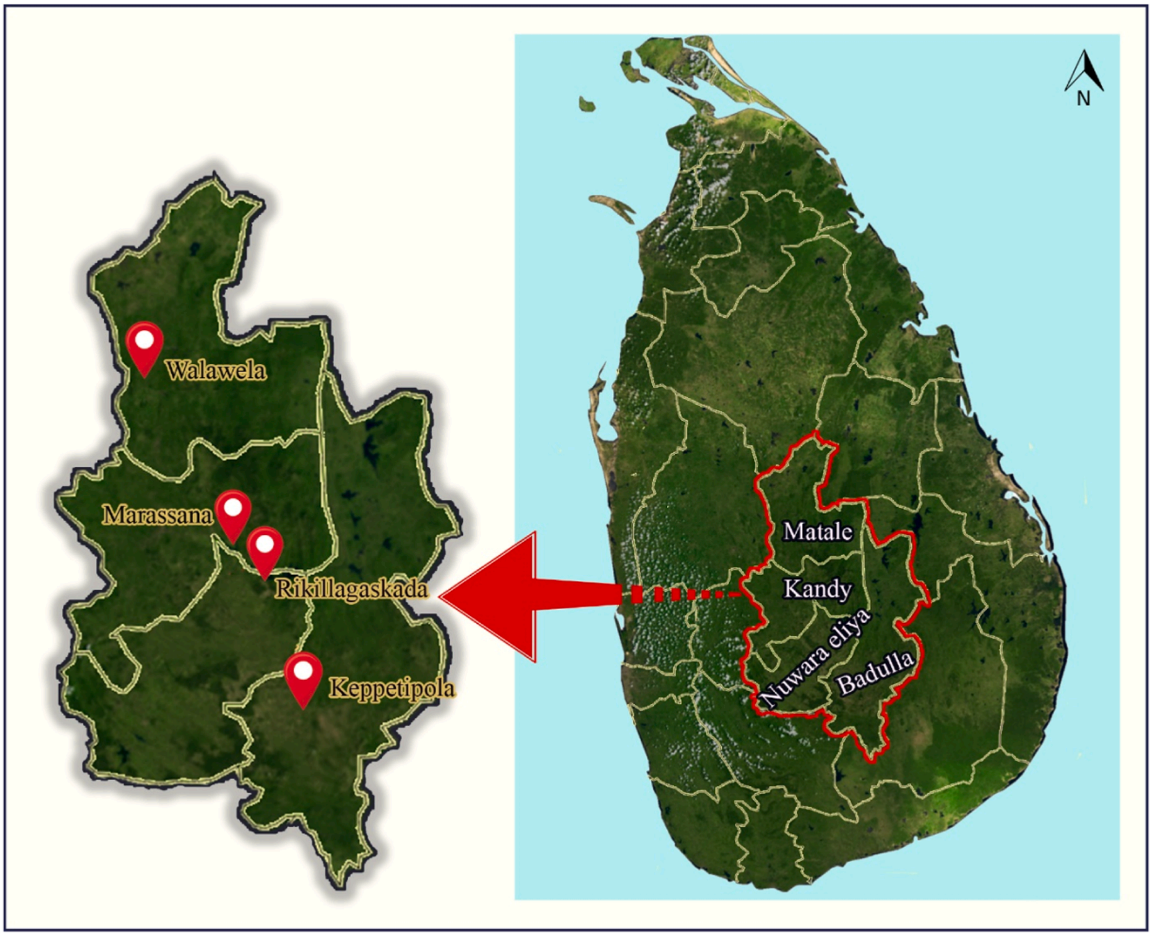

Fig. 2. Tomato sample selected distinct locations in Sri Lanka (Source: Authors' compilation).

these greenhouse locations, resulting in a total of 100 tomatoes directed towards the study.

These selected locations were situated within well-maintained greenhouses, each meticulously maintaining standardized conditions and agricultural protocols for tomato cultivation. These encompassed precise control of variables such as temperature, humidity, lighting, ventilation, irrigation management, nutrient and fertilizer delivery systems, as well as comprehensive pest and disease control measures (Table 1). It took 40 to 45 days from the flowering stage of tomatoes to reach the green mature stage. At this green tomato stage, they were harvested for analysis.

Preliminary study on tomatoes

As tomatoes undergo ripening, they exert an influence on environmental conditions, encompassing factors such as temperature, light intensity, and relative humidity [27,28]. To gain insights into the market environmental conditions, a preliminary study was conducted in significant market areas, including Colombo, Peliyagoda, Gampaha, Kandy, Dambulla, Anuradhapura, Galle, Kegalle, Trincomalee, Batticoloa, Mannar, and Jaffna in Sri Lanka (Fig 3A). This study involved the measurement of temperature, illumination levels, and relative humidity in 60 tomato-selling marketplaces across these regions (Fig 3B).

Table 1
Key Factors in Greenhouse Tomato Cultivation.

Condition	Levels
Plant Population	1 per 4 square feet
Growing Temperature	21 °C-27 °C
Relative Humidity	70 %
Irrigation for matured plant	2700 ml of water per plant per day
Light	80 % from outdoor light intensity
Nutrient solution	
pH	5.6–5.8
Nitrogen	50–200 ppm
Total dissolved solids	450–1600 ppm
Electroconductivity	0.6–2.2 mmhos

The relative humidity and temperature were measured using a humidity-temperature meter (RH390, MRCLAB, UK), and illumination was measured using a lux meter (LXP-10A, SONEL S.A., PL) at each marketplace. Environmental data at each marketplace were recorded for a duration of one year, and the average data was collected for the study. According to the study, the average ambient relative humidity, temperature, and illumination were 80 %, 29 °C, and 810 lux, respectively, and these average ambient conditions were used to prepare a tomato storage facility. The tomato storage facility is an artificially created environment that replicates the ambient conditions previously studied in the marketplaces, providing uniform conditions for the tomatoes without induce any physical damage to tomatoes. The stored tomatoes are also of good quality, without any physical damage or pest attack.

Development of the convolutional neural network

Image processing techniques are employed to discern tomatoes from other vegetables and fruits, assess their maturity levels, and estimate the number of days post-harvesting. This purpose is achieved through the utilization of a Convolutional Neural Network (CNN) models, with RGB values and shape serving as the key features.

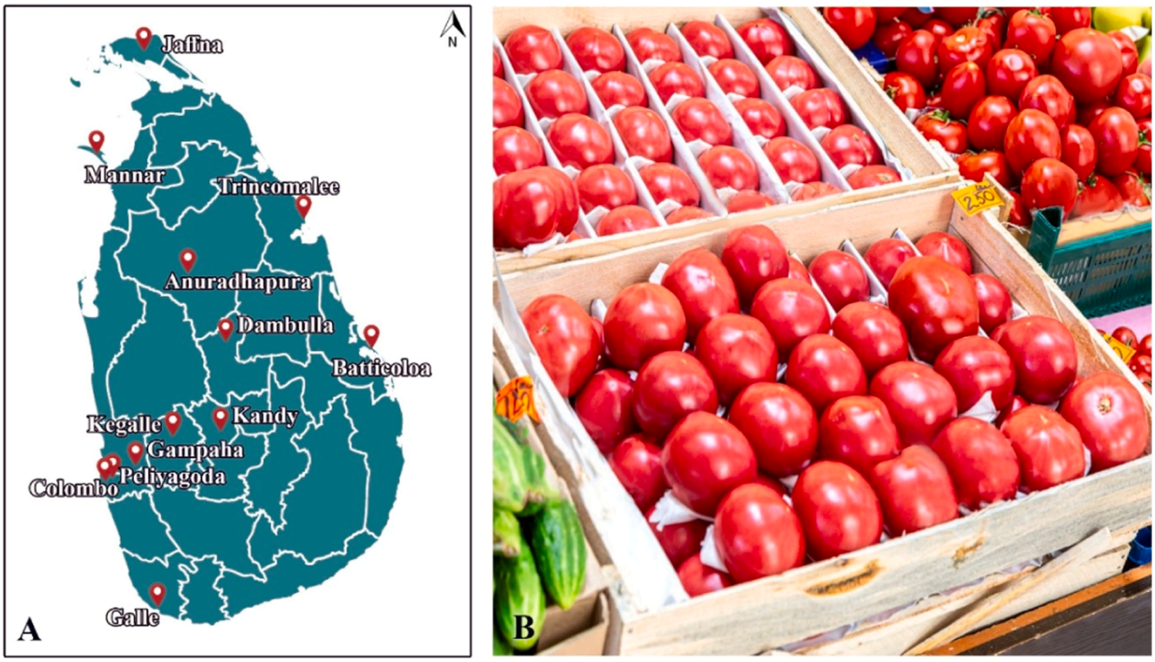

Fig. 3. Locations of the selected tomato marketplaces across Sri Lanka. (A) Selected tomato marketing places in Sri Lanka. (B) View of the tomato marketing place in Sri Lanka (Source: Authors' Compilation).

Data preparation for models (Creation of dataset)

Initially, images of tomatoes were captured using a 12 MP, f/2.4 camera from four sides of the tomato. The camera had a focal length of 4 mm, an exposure time of 1/20 second, and an ISO speed of 320. The photos had a resolution of 3024 × 4032. The sides were consistently maintained, and the capturing angles were not altered. Sides 1, 2, and the upper side were captured at a 90° angle, while images a 45° angle from upper side were captured at 45° (Fig 4A and Fig 4B). Photos of the tomatoes were taken daily at 22.30 p.m. Eastern Standard Time until they deteriorated. When capturing the tomatoes, consistent lighting conditions of 810 lux were maintained for all tomatoes.

In addition to determining the stage of tomato maturity, there was a need for a classification method to identify tomatoes suitable for consumption. Therefore, the IPGRI/IITA 1998 tomato classification method was employed to distinguish consumable tomatoes [29]. Only the excellent, good, and acceptable tomatoes were used to determine the stages of tomato maturity, as indicated in Table 2.

The approach of image processing techniques using CNN

In recent years, there's been a marked evolution in visual recognition techniques. Traditional classification methods have broadened their scope to adapt to larger-scale [30,31] scenarios, owing to the enhanced computational capabilities, notably the advent of powerful Graphics Processing Units. The efficacy of CNNs [32,33] has notably surpassed that of earlier approaches like the Bag-of-Visual-Words [34,35,36] and various compositional models [37,38]. The CNN serves as a structured model designed for comprehensive visual recognition tasks. Stemming from the premise that a sufficiently neuron-rich network can adapt to complex data distributions, neural networks in recent history have proven efficacious for rudimentary recognition tasks [37,39].

In addition to that, the realm of image processing, generally used the models of Support Vector Machines, K-Nearest Neighbors, and CNN for image classification process according to the accuracy level of each model [40]. According to this situation, the images have to undergone preprocessing via a range of distinct algorithms, enabling the assessment of the quantitative analytical effectiveness of diverse preprocessing techniques.

As image processing progresses across an expanding spectrum of dimensions, a continual improvement in the classification accuracy of defect images is consistently observed. However, it is noteworthy that the most pronounced enhancement in accuracy is observed in the case of the CNN algorithm [40].

A CNN consists of multiple layered structures. Within these layers, outputs from a preceding layer undergo convolution processes with a set of filters and are subsequently activated via a differentiable non-linear operation. Given this architecture, CNNs can be conceptualized as an integrated function, optimized through the back-propagation of error signals that emerge from disparities between the predicted outcomes and the supervisory input. In contemporary research, various methodologies, including ReLU (Rectified Linear Unit) activation function [32], Dropout [41], batch normalization [42], and Disturb Label [43], have been introduced to accelerate the convergence of CNNs and deter potential overfitting.

In the system, two CNN models have been developed based on the primary objective and the system's structural framework. Hence, in the preprocessing phase of the datasets for Model 1 and Model 2, several transformations were conducted on the image datasets to ensure their suitability for subsequent analysis. Initially, all images were resized to a consistent dimension of 128 × 128 pixels using bilinear interpolation. This resizing operation was essential for standardizing the dataset and facilitating uniform processing across all images.

After resizing, pixel values were subsequently normalized by dividing them by 255 to scale them to the range between 0 and 1. This normalization procedure effectively scaled the pixel intensities and ensured that the input data were appropriately standardized for subsequent modeling tasks.

The first objective is for the model to accurately identify an uploaded image, whether it is a tomato or any other type of vegetable. To achieve this, a CNN architecture with four primary layers was employed. Additionally, dense and dropout layers were included for the Model 1 (Fig 5).

Subsequently, another model was tasked with determining whether the tomatoes can be consumed and calculating the number of days elapsed after harvesting based on the tomato's maturity level.

Model 2 classifies the tomato into seven distinct categories as final classification of the system, including non-consumable and consumable six predefined maturity stages categorized as consumable: Green,

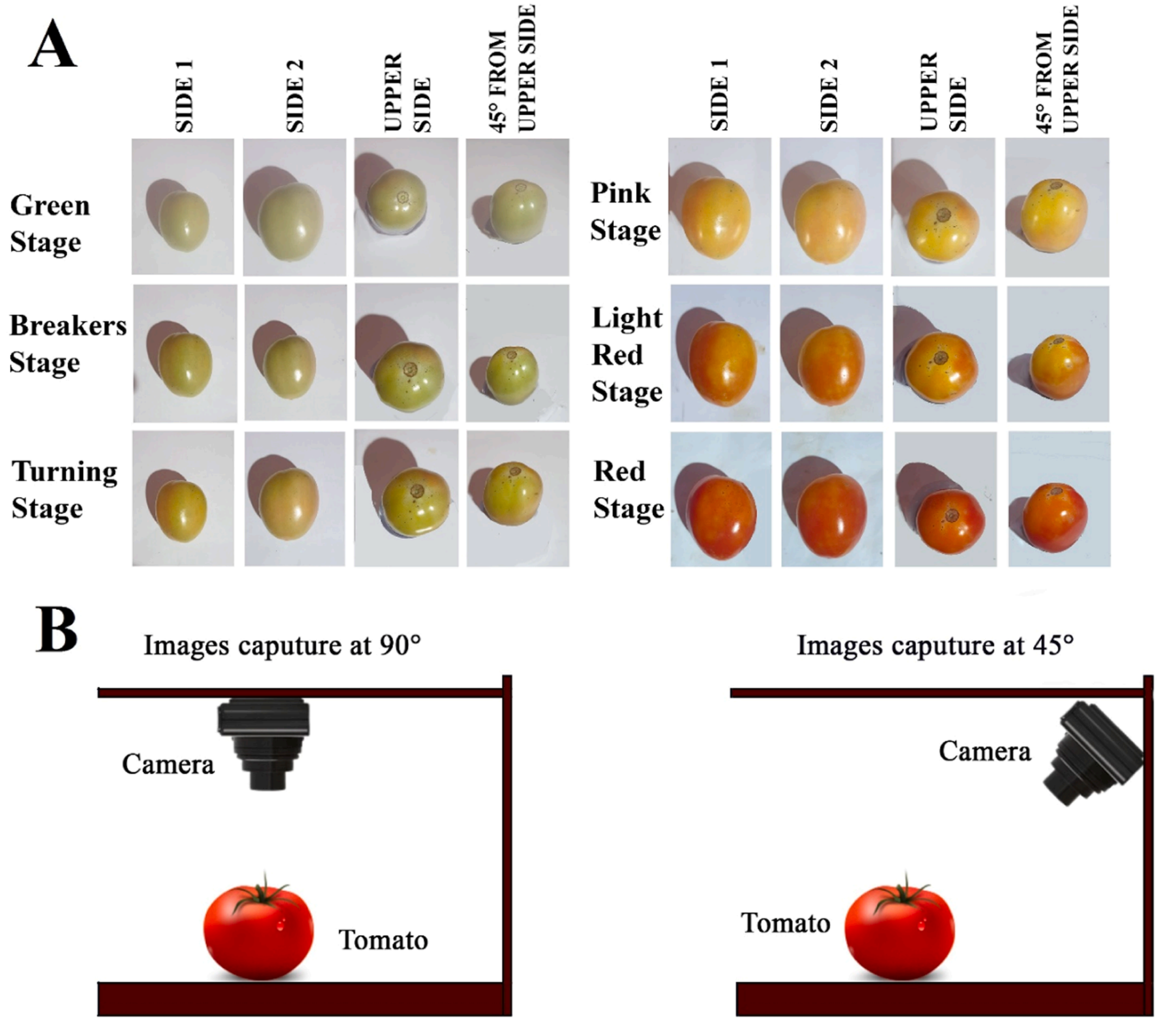

Fig. 4. Captured images of a tomato. (A) Captured tomatoes at different angles and stages at various times. (B) Method used to capture tomatoes. (Source: Authors' Compilation).

Table 2
Rating scales employed for assessing the overall freshness of tomato fruits [29].

Rating descriptor	Percent	Rating
Excellent	0	Overall appearance excellent; Fresh fruits with no major defects
Good	25	Overall appearance good, some fruits showing little or no sign of rottening/ blackening at the tips; few defects; acceptable firmness of fruits
Acceptable	50	Appearance of the limit of being acceptable; fruits flaccid; fruit tips brown/black; some with brown streaks; several/ many defects
Unacceptable	75	No freshness, fruits with black streaks and brown grooves
Poor	100	Musty odour, brown, slimy/decayed, many streaks and blackness on fruits

Non-tomato images were obtained from the "https://www.kaggle.com" open platform for dataset.

Breakers, Turning, Pink, Light Red, and Red. To achieve this, a CNN architecture with three primary layers is utilized. Furthermore, dense and dropout layers were incorporated for Model 2 as well (Fig 6).

These models utilize activation functions, including ReLU, sigmoid, and softmax, to enable this classification process. The entire computational framework was implemented using Python as the programming language, leveraging libraries such as TensorFlow, Matplotlib and numpy etc. The computational environment used for this study was Google Colab, providing the necessary resources for training and evaluation.

To achieve the primary objective of the system through Model 2, greenhouse-cultivated, pest-free tomatoes were used as the dataset to train the model. Additionally, harvested tomatoes were stored under standard market conditions to simulate real-world scenarios. The process flow diagram illustrating the entire study is depicted in Fig 7.

Activation functions

ReLU. In these two models, both the input and hidden layers, this study incorporated ReLU as the non-linear activation function [44,45]. While the ReLU is neither smoothed nor bounded, it has proven effective in applications that encompass a significant number of units, such as CNNs. ReLU is a piecewise linear function that outputs the input directly if it's positive; otherwise, it outputs zero. When employing Stochastic Gradient Descent with backpropagation of the errors for training, it's imperative to leverage a function that approximates linear behavior. Even though ReLU exhibits this linear-like behavior, it remains a nonlinear function, thus enabling the model to capture and learn intricate relationships present in the data [46,47].

Sigmoid activation function. The sigmoid function, also referred to as the logistic function, and exhibits several notable characteristics. it is

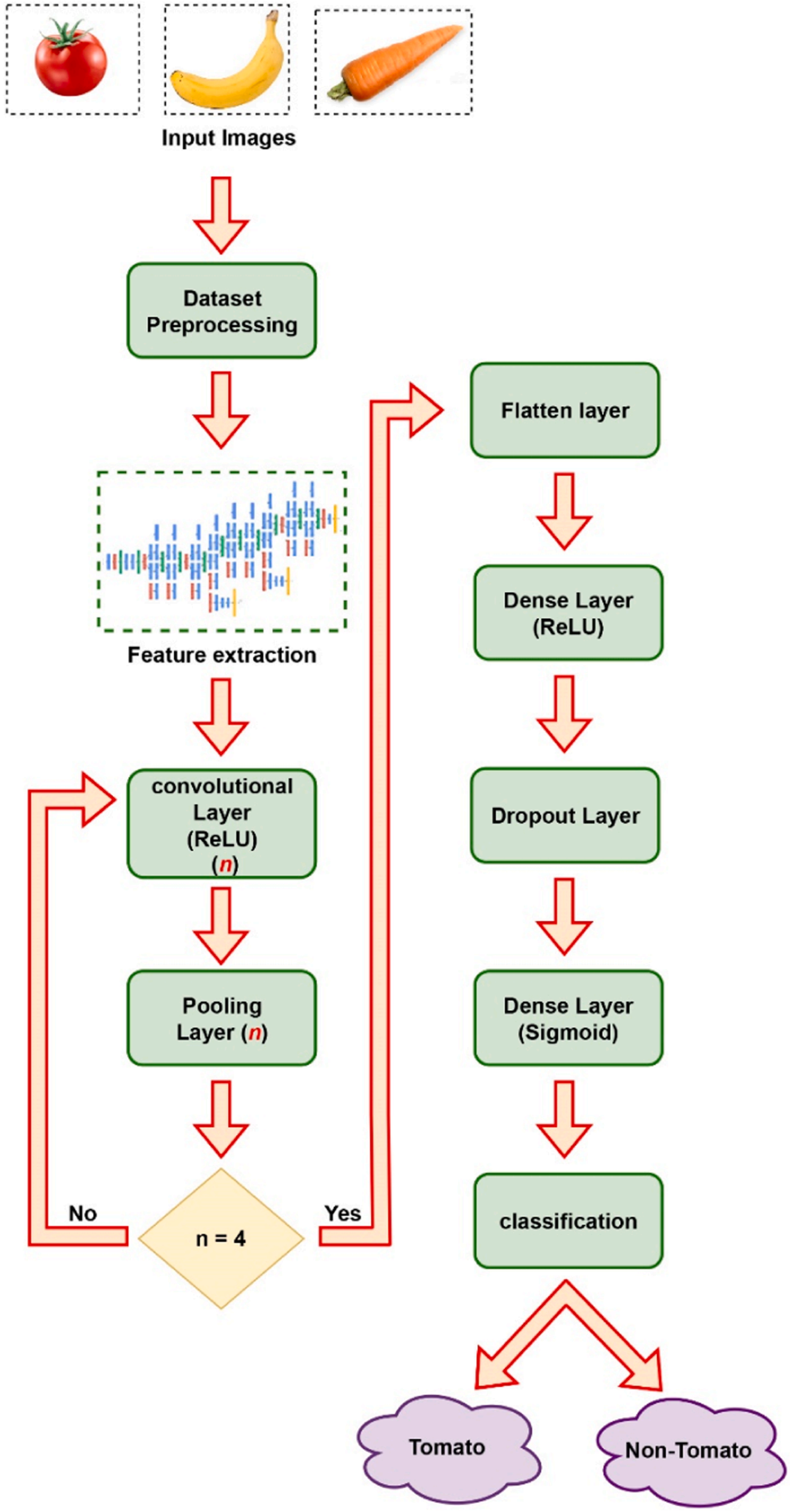

Fig. 5. Diagram of Model 1.

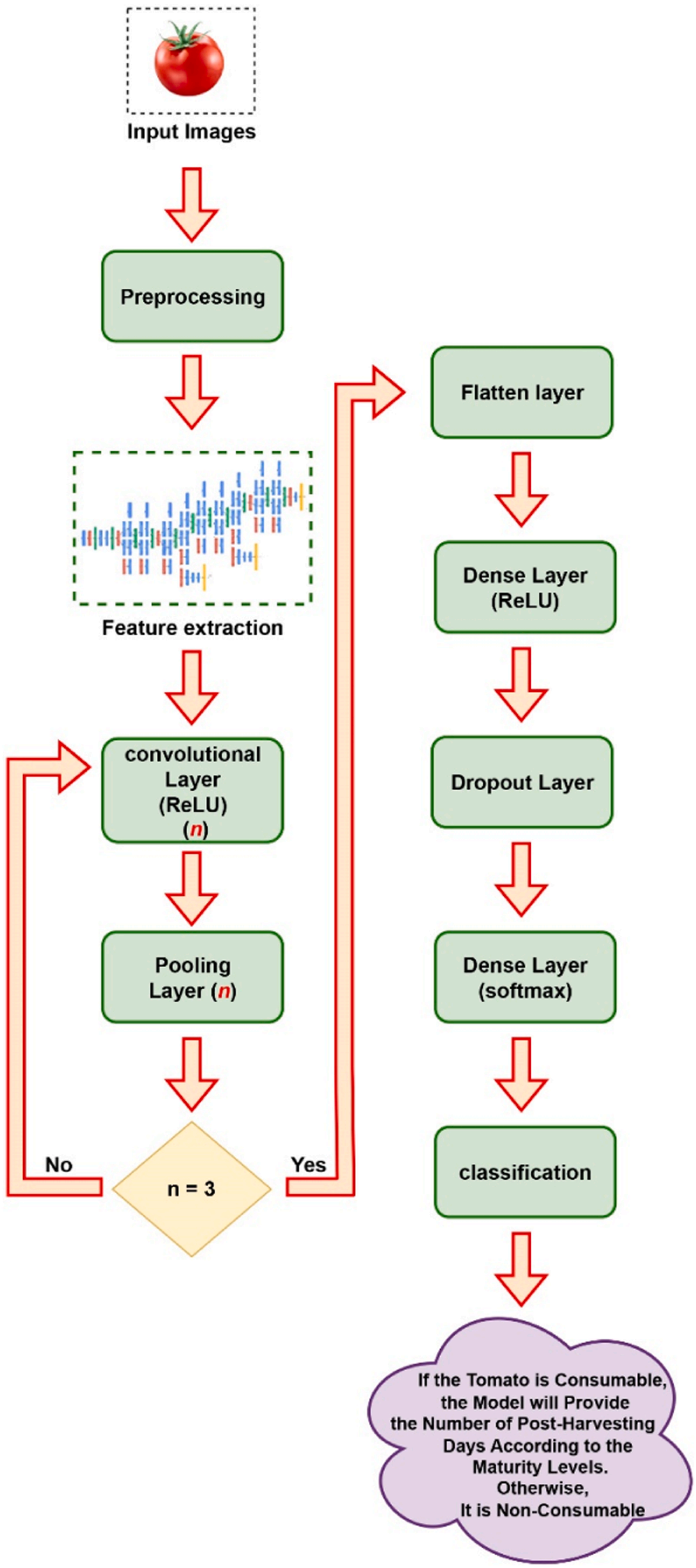

Fig. 6. Diagram of Model 2.

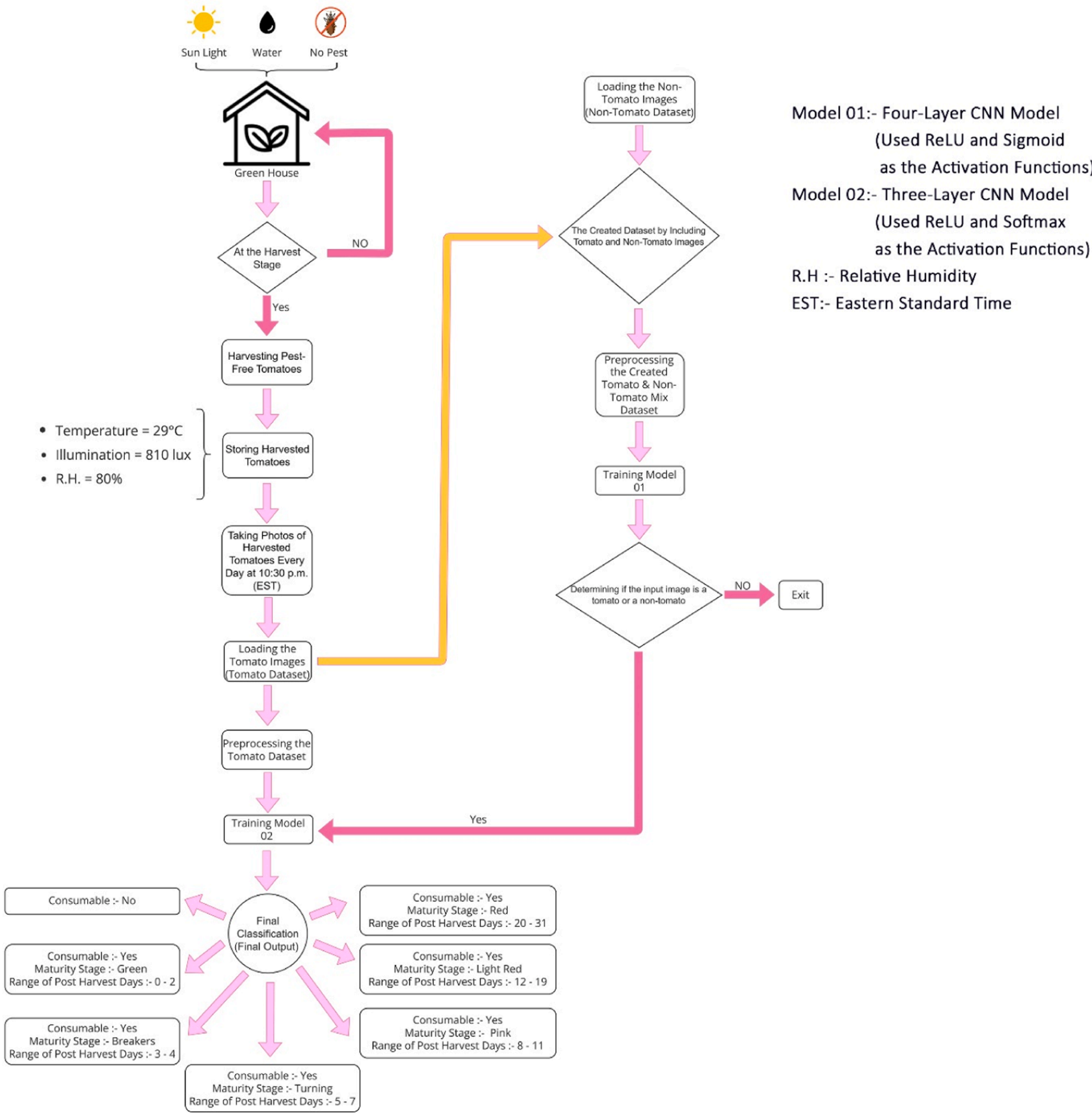

Fig. 7. Process flow diagram of the study.

continuous, monotonically increasing, and bounded. Specifically, the values produced by the sigmoid function lie between 0 and 1 [48–51]. For the purposes of our study, a value of "0″ to represent other vegetables and fruits (Fig 8A) and "1″ was assigned to denote tomatoes (Fig 8B) during the training phase in the first model. It aided us in the development of a high-performing model, enabling us to explore enhancements to the model and facilitating predictions on previously unseen data.

Softmax activation function. The softmax function constrains the output values within the range of 0 and 1, implying that the output can be contextualized as a probability distribution [52]. This characteristic means that the softmax function can be seen as an extension of the sigmoid function tailored for multiclass scenarios. Specifically, while the sigmoid function is well-suited for binary classification, the softmax function is adept at determining the probabilities of multiple classes concurrently [50,53,54]. Consequently, this activation function was employed to discern the non-consumable tomato and six distinct stages of maturity in second model.

Typically, the softmax function is utilized in the final layer of neural networks. A crucial observation to highlight is that the number of units in the softmax layer should correspond directly to the number of categories in the output layer, ensuring that each unit can represent the probability of a specific class.

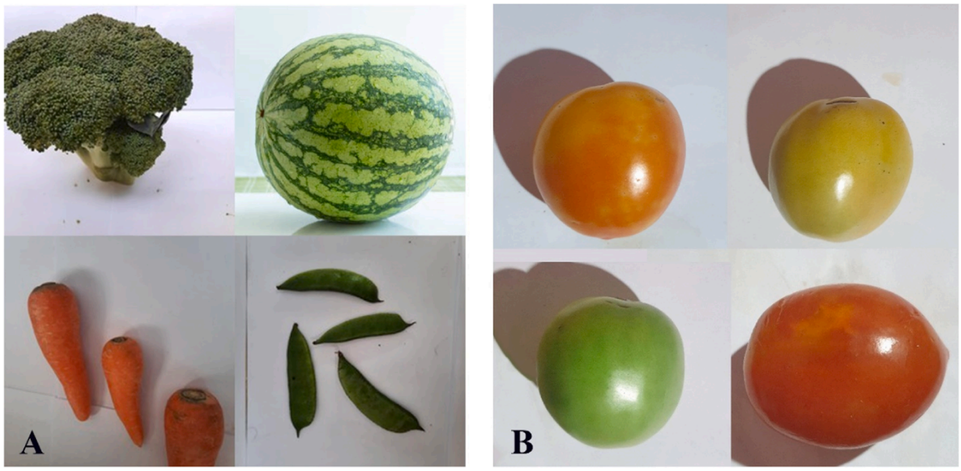

Fig. 8. Labeled data for sigmoid activation function. (A) Non Tomato (B) Tomato (Source: Authors' Composition).

Results and discussion

Graphically evaluation of model 1 and model 2

The accuracy plots for the two developed models depict the performance of each model throughout their training phases. These plots offer a visual representation of the model's accuracy on a specified dataset across successive epochs (Fig 9A and Fig 10A plots).

The loss plots for the two developed models provide a graphical representation of the evolution of the loss function throughout the training process. The loss function quantifies the discrepancy between predicted and actual values within a given dataset [55]. Monitoring the loss plot is crucial for comprehending and overseeing a model's training progression (Fig 9B and Fig 10B plots).

Fig 9A illustrates a plot depicting the training and validation accuracy of Model 1. The X-axis denotes the number of training epochs for model 1, each epoch signifying a complete traversal of the entire training dataset, ranging from 1 to 10. The Y-axis corresponds to model 1′s accuracy on both training and validation of the dataset. This metric reflects the model 1′s proficiency in classifying and predicting according to the two output classes, with values expressed between 0 and 1, representing the fraction of samples predicted accurately (Fig 9A).

Fig 9B illustrates a plot depicting the training and validation loss of Model 1. The X-axis represents each epoch within the entire training process. During each epoch, Model 1 is exposed to all the training examples exactly once, and its parameters are updated based on the calculated loss (Fig 9B).

Fig 10A illustrates a plot depicting the training and validation accuracy of Model 2. The X-axis represents the count of training epochs for Model 2, where each epoch signifies a full iteration through the entire training dataset, spanning from 1 to 10. The Y-axis represents the accuracy of Model 2 on both the training and validation datasets [56]. This metric assesses the ability of Model 2 to classify and predict data across seven output classes, with values ranging from 0 to 1, indicating the proportion of accurately predicted samples (Fig 10A).

Fig 10B illustrates a plot depicting the training and validation loss of Model 2. The X-axis enumerates each epoch throughout the comprehensive training phase. Within every epoch, Model 2 processes the entire set of training examples a single time, adjusting its parameters in accordance with the determined loss (Fig 10B).

Numerically evaluation of model 1 and model 2

This operational mechanism ensures that images classified as non-tomatoes do not proceed further in the system. Model 1 achieves an impressive training and validation accuracy of 99 %, while Model 2 exhibits a commendable training and validation accuracy of 99 % and 98 % respectively. The utilization of these widely recognized and established models significantly enhances the depth and credibility of our study, thus elevating its overall significance and reliability [57].

In model 1, the sigmoid function was employed as the activation function in the final layer, given the binary nature of the outputs in this

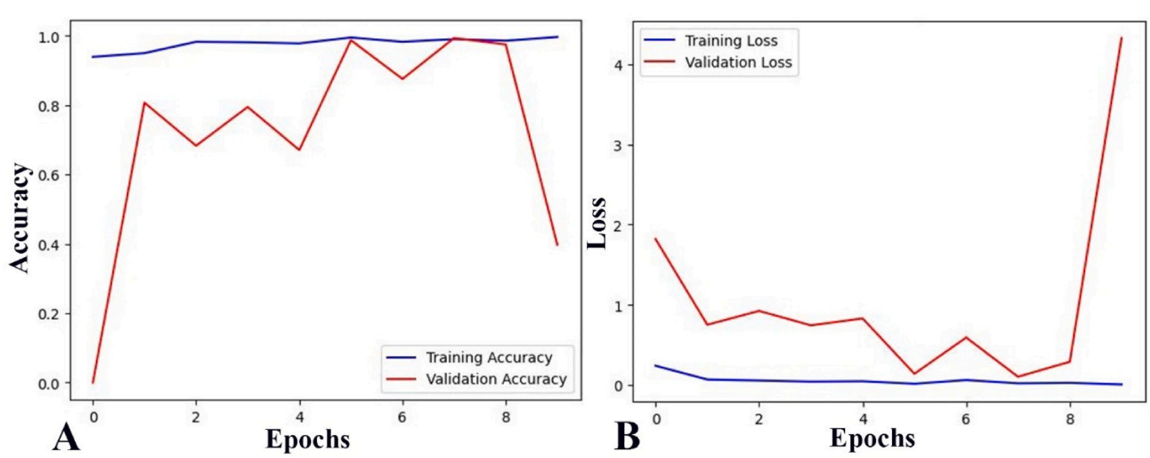

Fig. 9. Performance of Model 1. (A) Accuracy graph (B) Loss graph.

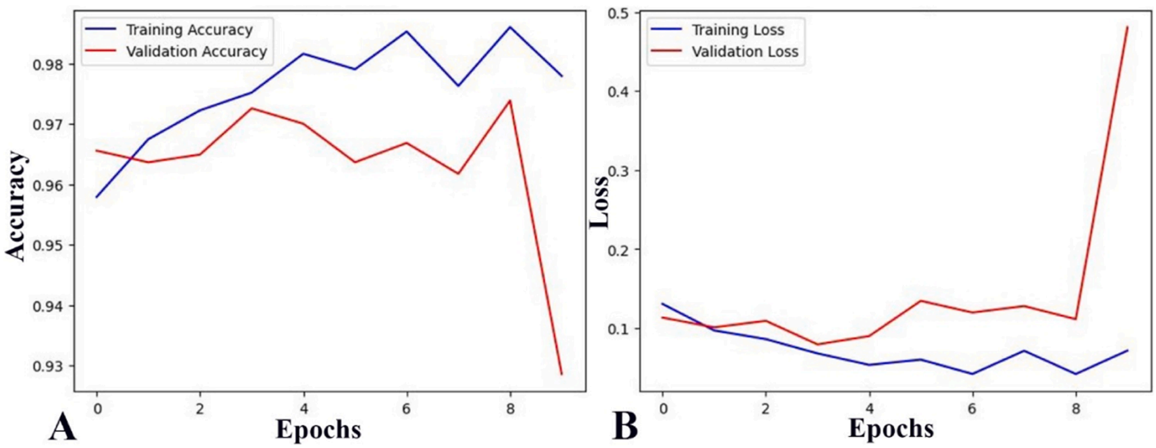

Fig. 10. Performance of Model 2. (A) Accuracy graph (B) Loss graph.

model. During the training process, 10 epochs were executed, resulting in superior validation accuracy and minimal validation loss achieved in the seven epoch (Table 3), with values of 0.99 and 0.11, respectively shown in Fig 9A and 9B.

In the model 2, the softmax function implemented as the activation function in the output layer, a choice driven by the presence of seven distinct output classes within this model. During training, 10 epochs were executed this model as well. Notably, by the conclusion of the eight epoch (Table 3), optimal performance metrics was observed, displaying validation accuracy of 0.98 and the validation loss of 0.11, respectively as shown in Fig. 10A and 10B.

The data collection process revealed an interesting observation regarding the number of days tomatoes existed in specific maturity stages. Remarkably, there appeared to be a consistent pattern across most tomatoes, indicating a nearly uniform duration in each particular stage. For instance, upon harvesting, tomatoes in the green stage were found to persist for approximately 2 days. As they transitioned to the breaker stage, this duration remained consistent at 2 days. Subsequently, during the turning stage, tomatoes exhibited an average existence of 3 days. Upon further transformation into the pink stage, this timeframe expanded to approximately 4 days. As tomatoes progressed into the light red stage, they displayed remarkable endurance, extending their presence for about 8 days. Finally, when reaching the red stage, tomatoes demonstrated the longest staying power, lingering for an average of 12 days (Table 4).

Table 3
The result of Model 1 and 2 during the training process for 5 to 10 epochs.

Epoch	Result of Model 1	Result of Model 2
Epoch 5/10	Training accuracy: 0.9783 Training loss: 0.0501 Validation accuracy: 0.9533 Validation loss: 0.432	Training accuracy: 0.9816 Training loss: 0.0534 Validation accuracy: 0.964 Validation loss: 0.1354
Epoch 6/10	Training accuracy: 0.9953 Training loss: 0.019 Validation accuracy: 0.8746 Validation loss: 0.0832	Training accuracy: 0.9895 Training loss: 0.04 Validation accuracy: 0.9662 Validation loss: 0.1245
Epoch 7/10	**Training accuracy: 0.9907** **Training loss: 0.0247** **Validation accuracy: 0.9938** **Validation loss: 0.1065**	Training accuracy: 0.9783 Training loss: 0.0819 Validation accuracy: 0.9668 Validation loss: 0.1197
Epoch 8/10	Training accuracy: 0.9829 Training loss: 0.0655 Validation accuracy: 0.9648 Validation loss: 0.325	**Training accuracy: 0.986** **Training loss: 0.0419** **Validation accuracy: 0.9769** **Validation loss: 0.1112**
Epoch 9/10	Training accuracy: 0.986 Training loss: 0.0306 Validation accuracy: 0.5752 Validation loss: 3.6432	Training accuracy: 0.9763 Training loss: 0.0713 Validation accuracy: 0.9377 Validation loss: 0.3279
Epoch 10/10	Training accuracy: 0.9969 Training loss: 0.0107 Validation accuracy: 0.3975 Validation loss: 4.3225	Training accuracy: 0.9779 Training loss: 0.0715 Validation accuracy: 0.9286 Validation loss: 0.4808

The culmination of these findings points to a significant revelation: tomatoes in the red stage tend to endure the longest before deterioration sets in. This discovery highlights the resilience of tomatoes at the red stage, suggesting a potential avenue for optimizing storage and utilization strategies to prolong their shelf life. Further research into the factors contributing to the extended vitality of red-stage tomatoes could unlock valuable insights for both farmers and consumers, ultimately enhancing the management of tomato quality and supply chains.

Tomatoes that do not align with any of the aforementioned categories are classified as non-consumable. According to the IPGRI/IITA 1998 tomato classification method, it includes Unacceptable and Poor conditions of tomatoes. Within the non-consumable category, two distinct subtypes emerge such as tomatoes deemed unfit for harvest and those that have succumbed to rotting [58]. This meticulous categorization scheme empowers a more nuanced understanding of tomato quality and suitability for consumption.

Conclusions

In this extensive research study, a journey was undertaken to explore and understand various aspects of tomato quality assessment. From the initial stages of capturing images of tomatoes in same environmental conditions to the intricacies of employing two distinct CNN models, the goal was to comprehensively evaluate tomato quality based on their external characteristics. One of the key highlights of this study was the careful selection of two CNN models, each tailored to specific requirements. This mechanism ensured that only valid tomato images progressed through the analysis process, saving computational resources and enhancing efficiency. To achieve this, Model 1 achieved training and validation accuracy level of 99 %. Meanwhile, Model 2 achieved training and validation accuracy level of 99 % and 98 % respectively, in detecting Padma tomato maturity stages and post-harvested day range under market conditions in Sri Lanka. Hence, the successful

Table 4
Days after harvesting based on the maturity stage.

Maturity Stage	Number of Existing Days of the maturity stage	Post-Harvest Days Range of the maturity stage
Green	2	0 – 2
Breakers	2	3 – 4
Turning	3	5 – 7
Pink	4	8 – 11
Light Red	8	12 – 19
Red	12	20 – 31

determination of post-harvest duration based on the six distinct maturity stages was achieved. In addition to that, the need for a robust operational mechanism to efficiently distinguish between tomatoes and non-tomatoes was addressed. In conclusion, this research underscores the importance of accurate and efficient tomato quality assessment. Through the strategic use of CNN models, the careful selection of activation functions, and the inclusion of a method like IPGRI/IITA 1998, our understanding of tomato quality has been advanced, and a robust framework for future research in this domain has been provided.

In light of our current findings, there are numerous prospective avenues for future research that hold promise. An expansion of the evaluation across diverse tomato varieties can provide a comprehensive understanding of their varying shelf lives. The specific duration that tomatoes can be retained before deterioration under the market conditions prevalent in Sri Lanka remains a crucial query. It is essential to delve deeper into the conditions that can prolong the storage life of tomatoes, potentially leading to reduced waste and increased economic benefits. Moreover, examining how storage conditions influence the deviation in the quality of tomatoes can shed light on optimal storage practices. These endeavors will not only bolster our current knowledge but also offer practical insights for stakeholders in the agricultural sector.

CRediT authorship contribution statement

Thilina Abekoon: Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Data curation. **Hirushan Sajindra:** Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **J . A . D . C . A . Jayakody:** Writing – original draft, Validation, Conceptualization. **E . R . J Samarakoon:** Writing – original draft, Validation. **Upaka Rathnayake:** Writing – review & editing, Validation, Supervision, Project administration, Conceptualization.

Declaration of 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.

Data availability

Data will be made available on request.

Funding

This research received no external funding.

References

[1] C.M. Rick, Tomato, Hybridizat. Crop Plants (2015) 669–680, https://doi.org/10.2135/1980.hybridizationofcrops.c48.
[2] P. Bhatia, N. Ashwath, T. Senaratna, D. Midmore, Tissue culture studies of tomato (lycopersicon esculentum), Plant Cell Tissue Organ Cult. 78 (1) (2004) 1–21, https://doi.org/10.1023/b:ticu.0000020430.08558.6e.
[3] M. Quinet, T. Angosto, F.J. Yuste-Lisbona, R. Blanchard-Gros, S. Bigot, J. P. Martinez, et al., Tomato fruit development and metabolism, Front. Plant Sci. 10 (2019), https://doi.org/10.3389/fpls.2019.01554.
[4] R. Martí, S. Roselló, J. Cebolla-Cornejo, Tomato as a source of carotenoids and polyphenols targeted to cancer prevention, Cancers 8 (6) (2016) 58, https://doi.org/10.3390/cancers8060058.
[5] A. Raiola, M.M. Rigano, R. Calafiore, L. Frusciante, A. Barone, Enhancing the health-promoting effects of tomato fruit for biofortified food, Mediators Inflamm. 2014 (2014) 1–16, https://doi.org/10.1155/2014/139873.
[6] Z. Liu, S. Alseekh, Y. Brotman, Y. Zheng, Z. Fei, D.M. Tieman, et al., Identification of a solanum pennellii chromosome 4 fruit flavor and nutritional quality-associated metabolite QTL, Front. Plant Sci. 7 (2016), https://doi.org/10.3389/fpls.2016.01671.
[7] Y. Li, H. Wang, Y. Zhang, C. Martin, Can the world's favorite fruit, tomato, provide an effective biosynthetic chassis for high-value metabolites? Plant Cell Rep. 37 (10) (2018) 1443–1450, https://doi.org/10.1007/s00299-018-2283-8.
[8] Food and Agriculture Organization Statistics [Internet]. 2021 [cited 2023 Aug 27]. Available from: https://www.fao.org/faostat/en/#data/QCL.
[9] Tomato grades and standards [Internet]. USDA; 1991 [cited 2023 Aug 27]. Available from: https://www.ams.usda.gov/grades-standards/tomato-grades-and-standards.
[10] Crop production program [Internet]. Sri Lanka Ministry of Agriculture; 2023 [cited 2023 Aug 27]. Available from: https://www.agrimin.gov.lk/web/index.php/statistics?lang=en.
[11] M. Ciudad-Mulero, J. Pinela, A.M. Carvalho, L. Barros, V. Fernández-Ruiz, I. C. Ferreira, et al., Bioaccessibility of macrominerals and trace elements from Tomato (Solanum lycopersicum L.) farmers' varieties, Foods 11 (13) (2022) 1968, https://doi.org/10.3390/foods11131968.
[12] G.G. Priyankara, C.L. Karunarathne, K.H. Sarananda, M. Ariyaratne, Effect of maturity stage on ripening and quality characters of four tomato (solanum lycopersicum L.) varieties of Sri Lanka, Trop. Agric. Res. 28 (4) (2017) 496, https://doi.org/10.4038/tar.v28i4.8249.
[13] S.A.M.C. Samarakoon, A. Balasuriya, R.G.A.S. Rajapaksha, W.A.R.T. Wickramara, Molecular detection and partial characterization of Tomato Yellow Leaf Curl Virus in Sri Lanka, Pak. J. Biol. Sci. 15 (18) (2012) 863–870, https://doi.org/10.3923/pjbs.2012.863.870.
[14] Hordi crop – tomato [Internet]. [cited 2023 Sept 1]. Available from: https://doa.gov.lk/hordi-crop-tomato/.
[15] G.F.G.N. Oussou, R. Sikirou, S.A. Afoha, M.E. Dossoumou, S.A. Boukari, F. A. Komlan, et al., Resistance assessment of tomato (Solanum lycopersicum L.) and GBOMA (Solanum macrocarpon L.) cultivars against bacterial wilt caused by Ralstonia Solanacearum in Benin, Pakistan J. Phytopathol. 32 (2) (2020), https://doi.org/10.33866/phytopathol.030.02.0610.
[16] M.E. Dossoumou, R. Sikirou, A. Adandonon, J. Gonroudobou, Baba-Moussa L. Tomato hybrid and local varieties screened for resistance to bacterial wilt caused by Ralstonia solanacearum under Screen House and field conditions, Am. J. Plant Sci. 12 (08) (2021) 1222–1235, https://doi.org/10.4236/ajps.2021.128085.
[17] D. Ireri, E. Belal, C. Okinda, N. Makange, C. Ji, A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing, Artif. Intell. Agric. 2 (2019) 28–37, https://doi.org/10.1016/j.aiia.2019.06.001.
[18] V. Sahni, S. Srivastava, R. Khan, Modelling techniques to improve the quality of food using artificial intelligence, J. Food Qual. 2021 (2021) 1–10, https://doi.org/10.1155/2021/2140010.
[19] S. Laykin, V. Alchanatis, E. Fallik, Y. Edan, Image–processing algorithms for tomato classification, Trans. ASAE 45 (3) (2002), https://doi.org/10.13031/2013.8838.
[20] M. Haggag, S. Abdelhay, A. Mecheter, S. Gowid, F. Musharavati, S. Ghani, An intelligent hybrid experimental-based deep learning algorithm for Tomato-sorting controllers, IEEE Access 7 (2019) 106890–106898, https://doi.org/10.1109/access.2019.2932730.
[21] Y.H. Cheng, C.Y. Tseng, D-M Nguyen, Y.D. Lin, Yolov4-driven appearance grading filing mechanism: toward a high-accuracy tomato grading model through a deep-learning framework, Mathematics 10 (18) (2022) 3398, https://doi.org/10.3390/math10183398.
[22] B. Benmouna, R. Pourdarbani, S. Sabzi, R. Fernandez-Beltran, G. García-Mateos, J. M. Molina-Martínez, Comparison of classic classifiers, metaheuristic algorithms and convolutional neural networks in hyperspectral classification of nitrogen treatment in tomato leaves, Remote Sens. 14 (24) (2022) 6366, https://doi.org/10.3390/rs14246366.
[23] M. Kharbach, M. Alaoui Mansouri, M. Taabouz, H Yu, Current application of advancing spectroscopy techniques in food analysis: data Handling with chemometric approaches, Foods 12 (14) (2023) 2753, https://doi.org/10.3390/foods12142753.
[24] S. Wolfert, L. Ge, C. Verdouw, M.J. Bogaardt, Big Data in smart farming – a review, Agric. Syst. 153 (2017) 69–80, https://doi.org/10.1016/j.agsy.2017.01.023.
[25] E. Borràs, J. Ferré, R. Boqué, M. Mestres, L. Aceña, O. Busto, Data fusion methodologies for food and beverage authentication and quality assessment – a review, Anal. Chim. Acta 891 (2015) 1–14, https://doi.org/10.1016/j.aca.2015.04.042.
[26] G. Mahajan, K.G. Singh, Response of greenhouse tomato to irrigation and fertigation, Agric. Water Manage. 84 (1–2) (2006) 202–206, https://doi.org/10.1016/j.agwat.2006.03.003.
[27] E. Abiso, N. Satheesh, A. Hailu, Effect of storage methods and ripening stages on postharvest quality of tomato (Lycopersicom esculentum mill) cv. chali, Ann. Food Sci. Technol. (2015) 127–137.
[28] G. Mahajan, K.G. Singh, Response of greenhouse tomato to irrigation and fertigation, Agric. Water Manage. 84 (1–2) (2006) 202–206, https://doi.org/10.1016/j.agwat.2006.03.003.
[29] L.A. Babatola, D.O. Ojo, O.I. Lawal, Effect of storage conditions on tomato (Lycopersicon esculentum mill.) quality and shelf life, J. Biolog. Sci. 8 (2) (2008) 490–493, https://doi.org/10.3923/jbs.2008.490.493.
[30] J. Deng, W. Dong, R. Socher, L-J Li, K. Li, L. Fei-Fei, ImageNet: a large-scale hierarchical image database. Computer Vision and Pattern Recognition, IEEE Computer Society Conference, Miami, 2009.
[31] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, A. Oliva, Learning deep features for scene recognition using places database. Advances in Neural Information Processing Systems 27, The MIT Press, Montreal, 2014.

[32] A. Krizhevsky, I. Sutskever, G. Hinton, Learning deep features for scene recognition using places database, in: Advances in Neural Information Processing Systems 25, Montreal: The MIT Press, 2012.
[33] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: Proceedings of the International Conference on Learning Representations. Banff, 2014.
[34] G. Csurka, C. Dance, L. Fan, J. Willamowski, C. Bray, Visual categorization with bags of keypoints, in: Proceedings of the European Conference on Computer Vision, Prague, 2004.
[35] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, Y. Gong, Locality-Constrained Linear Coding for Image Classification. Computer Vision and Pattern Recognition, IEEE Computer Society Conference, San Francisco, 2010.
[36] F. Perronnin, J. Sanchez, T. Mensink, Improving the fisher kernel for large-scale image classification, in: Proceedings of the European Conference on Computer Vision, Heraklion, 2010.
[37] L. Xie, A. Yuille, Genetic CNN, in: Proceedings of the International Conference on Computer Vision, Venice, 2017.
[38] P.F. Felzenszwalb, R.B. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part-based models, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1627–1645, https://doi.org/10.1109/tpami.2009.167.
[39] Y. LeCun, J. Denker, D. Henderson, R. Howard, W. Hubbard, L. Jackel, Handwritten digit recognition with a back-propagation network. advances in neural information processing systems. Advances in Neural Information Processing Systems 2, The MIT Press, Colorado, 1989.
[40] H. Wu, Z. Zhou, Using convolution neural network for defective image classification of industrial components, Mobile Info. Syst. 2021 (2021) 1–8, https://doi.org/10.1155/2021/9092589.
[41] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn Res. 15 (2014) 1929–1958.
[42] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in: Proceedings of the International Conference on Machine Learning, Lille, 2015.
[43] L. Xie1, J. Wang, Z. Wei, M. Wang, Q. Tian, DisturbLabel: regularizing CNN on the loss layer, in: Proceedings of the International Conference on Computer Vision, Roorkee, 2016.
[44] K. Narayana Rao, K. Venkata Rao, P.V.G.D. PR, A hybrid intrusion detection system based on sparse autoencoder and deep neural network, Comput. Commun. 180 (2021) 77–88, https://doi.org/10.1016/j.comcom.2021.08.026.
[45] Q. Zhao, X. Chen, Y. Zhang, M. Sha, Z. Yang, W. Lin, et al., Synthesizing relu neural networks with two hidden layers as barrier certificates for Hybrid Systems, in: Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control, 2021, https://doi.org/10.1145/3447928.3456638.
[46] Agarap A.F.M. [Internet]. 2019 [cited 2023 Sept 2]. Available from: https://arxiv.org/pdf/1803.08375.pdf.
[47] M.R. Lakin, Design and simulation of a multilayer chemical neural network that learns via backpropagation, Artif. Life 29 (3) (2023) 308–335, https://doi.org/10.1162/artl_a_00405.
[48] S.S. Liew, M. Khalil-Hani, R. Bakhteri, Bounded activation functions for enhanced training stability of deep neural networks on visual pattern recognition problems, Neurocomputing 216 (2016) 718–734, https://doi.org/10.1016/j.neucom.2016.08.037.
[49] M. Kull, T.M. Silva Filho, P Flach, Beyond sigmoids: how to obtain well-calibrated probabilities from binary classifiers with beta calibration, Electron J. Stat. 11 (2) (2017), https://doi.org/10.1214/17-ejs1338si.
[50] M.A. Mercioni, S. Holban, The most used activation functions: classic versus current, in: Proceedings of the International Conference on Development and Application Systems, Suceava, IEEE, 2020.
[51] H. Sajindra, T. Abekoon, E.M. Wimalasiri, D. Mehta, U. Rathnayake, An artificial neural network for predicting groundnut yield using climatic data, Agri Eng. 5 (4) (2023) 1713–1736, https://doi.org/10.3390/agriengineering5040106. Sept 30.
[52] Y. Zhang, J. Wang, X. Zhang, Learning sentiment sentence representation with multiview attention model, Inf. Sci. 571 (2021) 459–474, https://doi.org/10.1016/j.ins.2021.05.044.
[53] H. Sajindra, T. Abekoon, J.A.D.C.A. Jayakody, U. Rathnayake, A novel deep learning model to predict the soil nutrient levels (N, P, and K) in cabbage cultivation, Smart Agric. Technol. 7 (2024 Mar) 100395, https://doi.org/10.1016/j.atech.2023.100395.
[54] T.K. Paradarami, N.D. Bastian, J.L. Wightman, A hybrid recommender system using artificial neural networks, Expert Syst. Appl. 83 (2017) 300–313, https://doi.org/10.1016/j.eswa.2017.04.046.
[55] M. Shivashankar, M. Pandey, M. Zakwan, Estimation of settling velocity using generalized reduced gradient (GRG) and hybrid generalized reduced gradient–genetic algorithm (Hybrid Grg-GA), Acta Geophys. 70 (5) (2022 Jan 4) 2487–2497, https://doi.org/10.1007/s11600-021-00706-2.
[56] M. Pandey, M. Jamei, I. Ahmadianfar, M. Karbasi, A.S. Lodhi, X. Chu, Assessment of scouring around spur dike in cohesive sediment mixtures: a comparative study on three rigorous machine learning models, J. Hydrol. 606 (2022 Mar) 127330, https://doi.org/10.1016/j.jhydrol.2021.127330.
[57] M. Pandey, M. Zakwan, P.K. Sharma, Z. Ahmad, Multiple linear regression and genetic algorithm approaches to predict temporal scour depth near circular pier in non-cohesive sediment, ISH J. Hydraulic Eng. (2018 Apr 19) 1–8, https://doi.org/10.1080/09715010.2018.1457455.
[58] N.S. Bhatkar, S.S. Shirkole, A.S. Mujumdar, B.N. Thorat, Drying of tomatoes and tomato processing waste: a critical review of the quality aspects, Drying Technol. 39 (11) (2021) 1720–1744, https://doi.org/10.1080/07373937.2021.1910832.