

A Transfer Learning-based Approach to Predict the Shelf life of Fruit

Varsha Bhole^[1,A], Arun Kumar^[1,B]

^[1] Department of Computer Science and Engineering, Sir Padampat Singhania University, Udaipur, 313601 India.

vybhole@acpce.ac.in^[A], arun.kumar@spsu.ac.in^[B]

Abstract Shelf-life prediction for fruits based on the visual inspection and with RGB imaging through external features becomes more pervasive in agriculture and food business. In the proposed architecture, to enhance the accuracy with low computational costs we focus on two challenging tasks of shelf life (remaining useful life) prediction: 1) detecting the intrinsic features like internal defects, bruises, texture, and color of the fruits; and 2) classification of fruits according to their remaining useful life. To accomplish these tasks, we use the thermal imaging technique as a baseline which is used as non-destructive approach to find the intrinsic values of fruits in terms of temperature parameter. Further to improve the classification tasks, we combine it with a transfer learning approach to forecast the shelf life of fruits. For this study, we have chosen ‘Kesar’ (*Mangifera Indica* Linn cv. Kesar) mangoes and for the purpose of classification, our designed dataset images are categorized into 19 classes viz. RUL-1 (Remaining Useful Life-1) to RUL-18 (Remaining Useful Life-18) and No-Life as after harvesting, the storage span of ‘Kesar’ is near about 19 days. A comparative analysis using SqueezeNet, ShuffleNet, and MobileNetv2 (which are prominent CNN based lightweight models) has been performed in this study. The empirical results show a highest achievable accuracy of 98.15±0.44% with an almost a double speedup in training the entire process by using thermal images.

Keywords: Thermal Imaging, Deep Learning, Transfer Learning, ShuffleNet, Pre-trained models, Shelf life.

1 Introduction

Globally food wastage is a serious issue and it might become worst in near future. UN report shows that out of the total wasted food, every year around 14% of food produced is lost between harvest and retail, and \$181.2 crore is shared by fruit waste as per the report by Emerson Climate Technologies, India (2013). Further, some studies indicated that out of the overall fruit production, 40% (worth approximately \$8.3 billion) is simply wasted [1] whereas studies in the Netherlands have revealed that the wastage costs are around € 4.4 billion (US\$ 4.9 billion) [2] and also in Australia, nearly a AU\$ 1.1 billion (US\$ 0.84 billion) worth of fruits and vegetables is wasted [3]. One of the major causes is strict guidelines for accepting saleable fruits by supply centers. The agricultural supply network faces these wastage issues but the retail and customer sector face it on a higher level. So, to strengthen the economic viability and to escalate marketability, accurate shelf life evaluation of fruits is essential.

The detection of remaining useful life of fruit is still influenced by visual inspection, which also creates major difficulties. When the fruits are ready for harvesting, the maturity level is not identical for all the fruits which are collected from the same or different farms. So, it is essential to segregate the fruits according to their shelf life for transportation to various places. The prime task in fruit business is to segregate the fruits on the basis of color, size, texture, maturity rate of the fruits. This fruit segregation process reduces the transportation and cold storage problems, which are universal issues, but are more prevalent in South and Southeast Asia, followed by Africa, Europe, North America, and Oceania [4]. But with the use of thermal imaging technique combined with a deep

learning-based classification approach, by predicting shelf life presents an opportunity that benefits in reducing food wastage and will be helpful for replacement of physical measurements or human assistance.

Fruit maturity, an important factor to judge the shelf life is basically the opinion of an individual's perception based on their expertise which is subjective in nature. The methods to evaluate the fruit shelf life could be distinguished as objective and subjective methods with its pros and cons. Objective techniques are based on the sample's features; however, subjective ones are concerned with individual's perception. The objective approach is categorized as - destructive and non-destructive. Many of the destructive techniques uses less quantity of samples for examination but further these samples cannot be reused. The drawback of the destructive method is that the features which have been calculated from the selected sample may not always closely relate to the bulk from where the samples have been drawn, because of this there might be large variations in the result [5]. The statistical feature like linear regression and correlation coefficient may vary in the case of observed sample from that of the bulk from which the sample has been picked up [6]. On the contrary, samples remain intact, unhurt, and safe for consumers even post-testing in non-destructive techniques. Usually, destructive methods are performed at the laboratory level using some chemical analysis. Through the destructive approach one can measure moisture content, pH factor etc., to determine the shelf life of fruits, but it is not economical and feasible to a layman, and this may not be a handy work for all.

But in recent years, varieties of non-destructive techniques like computer vision, x-ray imaging, hyperspectral imaging, visible imaging, thermal imaging, etc. [7-9] have come up in the agricultural segment where utilization of thermal imaging is at the budding stage. So in this research, we introduce the deep thermal imaging technique to determine the shelf life of fruit, specifically mango fruit. This deep thermal imaging technique would be objective and cost-effective and will help to evaluate the exact shelf life which will make less the food wastage by segregating excessively ripe fruits which may be sold at the local market. Further, by using deep learning classification the margins for marketable fruits will be protracted. Then, ripeness standards would be evaluated and that will be beneficial to raise the returns. We have chosen the mango for a study. As mango fruit is locally available in bulk, and at the same time the present research was initiated during this season, therefore the mango fruits had been adopted and had become an obvious choice. The mango fruit is the most cultivated fruit in our local region and globally India ranks first in mango production; the mango fruit has a huge opportunity in the export market as well [10].

The application of our research is very beneficial for farmers, retailers, consumers as well as for micro-business related agro-industrial companies since it swiftly identifies the maturity and shelf life of fruits efficiently. The practical merit of this particular research is advantageous for fruit processing businesses and pack houses which require rapid, non-destructive, and easy to handle quality recognition system for prediction of shelf-life of fruits. This attempt is mainly aimed at providing access to technology innovations to improve the availability of fruits by increasing their production and prevention of postharvest losses which in turn to improve socio-economic life. The objective of the current work is to find the intrinsic features without destruction of the fruit to develop the accurate shelf-life prediction system with minimum resources as well as minimizing time complexity using thermal imaging via transfer learning through fine-tuned pre-trained MobileNetv2, ShuffleNet and SqueezeNet models.

The main contributions in the paper are:

1. To demonstrate the effective usage of thermal imaging techniques in finding the intrinsic features like internal defects, bruises, texture, and color.
2. To evaluate the performance of CNN-based lightweight models for feature learning.
3. To find out a better model to improve the classification tasks.

In this paper, the related work discusses in section 2. Section 3 covers the data collection and methodology; section 4 takes on results and discussion followed by section 5 for conclusion.

2 Related Work

The very first remarkable study has been performed using computer vision regarding grading in 1991, where an image processing technique via CCD camera to inspect quality of products on conveyor belts through 3D Look Up Table has been tried [11]. After that many researchers have contributed in assessing the quality of fruits in the last few decades including maturity using RGB imaging [12-23]. However, limited work has been implemented using thermal imaging on very few fruits like apple [24-27], citrus [28], peach [29], mango [30, 31, 32], blueberries [33], dates [34]. Recently, some researchers have proposed the applications in the agricultural domain related to computer vision with thermal imaging technique like defect detection and classification of fruits [7, 8, 33, 35], grading of fruits through maturity [30, 31, 34] and identifying the shell life [32].

Besides the above-mentioned work, the colossal research has been carried on the post-harvest physiology of the fruits. Since a majority of biochemical events during fruit ripening would seem to be genetically programmed; the ripening process provides a unique opportunity to study the senescence process systematically. In order to comprehend ripening phenomenon of mango, various observations have been noted based on the alterations which happen in physiochemical and biochemical level [36-38].

Beyond the mentioned literature, very little work has been done to estimate the shelf life of fruit. In [39], Near Infrared Spectroscopy was employed for analysing the quality of apple at various phases of shelf life and gained the accuracy up to 100%. In [40], stiffness of *Malus Asiatica Nakai* fruit was detected within its storage span using fast and non-destructive methodology such as Visible/ Near-Infrared Spectroscopy (VIS/NIR spectrometer). To find the fruit firmness, variations with dissipation of pectin is witnessed. Amongst three techniques (PLS, PCR and LS-SVM) used, the optimum one was the PLS technique at 1st and 4th day of shelf life which was evaluated by Standard Normalized Variation (SNV) with the correlation coefficients of prediction (Rp) of 0.7999 and 0.7494, and the Root-Mean-Square Errors of Prediction (RMSEP) of 0.7994 and 0.5856, respectively, which proposes a great association within the spectrum and stiffness of *Malus Asiatica Nakai*. But, 7th day of shelf life, weight reduction was more and quality has been degraded abruptly. The authors of [16, 17] have worked on five varieties of mangoes (Kumrapali, Amrapali, Sori, Langra and Himsagar) which have been sorted with 4 categories viz. poor (G1), medium (G2), good (G3) and very good (G4) dependent on maturity value and quality through the fuzzy rule-based algorithms. Maturity levels are identified using Gaussian Mixture Model (GMM) and quality was analysed using size and defects of fruits. Further, authors of [41] used the recursive feature elimination (RFE) method along with SVM-based classifiers for feature selection (13) amongst the initial 27 features and achieved the predictive classification accuracy up to 96% which determines the ripeness level of these mangoes with 4 classes for storage period remaining like M1 (12 days), M2 (within 9 to 12), M3 (within 5 to 8), M4 (within 1 to 4). In [42] machine learning techniques were utilized for clustering and classification to foresee banana maturity stages and shelf life. In the same work, images of 46 bananas have been captured daily for fourteen days and then subjected to color feature extraction. The resulting images have been clustered using k-means clustering along with seven different machine learning classification algorithms for the classification of banana shelf-life using external features. Amongst all, Decision Tree Classifier has achieved the highest accuracy around 52%.

In the artificial intelligence research field, deep learning (DL) with CNN comes as a handy tool [30, 35]. But, deep learning study and training a CNN from scratch requires good quality as well as large amount of data. For shelf-life evaluation, no specific dataset exist which can be used directly and creating a high-quality and huge dataset is a very lengthy and tedious task [43]. So, we have balanced tradeoff amongst a big but low-quality dataset as well as a small but good-quality dataset with transfer learning technique which has been found successful in the agricultural field with small dataset [30]. Transfer learning (TL) has a natural learning ability where one or more source data is transferred and utilized to enhance the learning of related new data; that is, it uses the pre-trained network, modify it, and re-train it based on new target data without start from scratch. The major steps involved in building a computational model using TL are: preparing a model and getting it trained on a large bench marked dataset; then follows the removal of the last layer i.e. "Loss output", with other working layers and finally the model is trained over a small dataset [43].

It is clear from the above survey that varied techniques have been employed for estimating shelf life of fruits but particularly artificial intelligence method like deep learning based transfer learning technique including thermal imaging has not come into picture yet from this perspective. In the current work, with an eye to study the maturity of mangoes during entire shelf life, the non-destructive method like deep thermal imaging has been adopted, for the first time in this research work to find their intrinsic features throughout the storage span according to the best of our knowledge.

3 Materials and Methods

The purpose of current study is to predict shelf-life of mangoes using state-of-the-art non-destructive approach, namely thermal imaging with three pre-trained models via transfer learning techniques. It has been ascertained that transfer learning requires small image datasets for training to gain high accuracy [30, 35]. Within this research, data collection has been pursued, as is explicated in Sect. 2.1. Thenceforth, in Sect. 2.2.1, process of proposed shelf-life prediction is presented. Sect. 2.2.2 provides an overview of the transfer learning-based pre-trained CNN (Convolutional Neural Network) architectures, namely SqueezeNet, ShuffleNet, and MobileNetV2. Subsequently, in Sect. 2.2.3, details of proposed ShuffleNet architecture is explained.

3.1 Sample collection and preparation

'Kesar' mango cultivar has been used in current work. These mangoes have been brought from three different places of India because of its varied size and quality. Every day, we observed 41 mangoes and photographed RGB as well as thermal images of the same. The videos and images have been axiomatically captured by regular smartphone camera and Seek Thermal camera from 360° as proposed by [7]. The captured RGB and thermal images have a resolution in pixels of 2322 x 4128 and 720 x 1280 respectively. This experimentation was repeated for around 18-19 days. The background, time to capture photographs, lighting conditions, and distance between camera and fruit were retained same during these days which is required for RGB imaging but not obligatory for thermal imaging. Nearby 10000 images of dataset have been produced for the current work. With the availability of remaining useful life, the images have been categorized into 19 classes' viz. RUL-1 (Remaining Useful Life-1) to RUL-18 (Remaining Useful Life-18) and No-Life. The Day_1 captured images were stored in class RUL_18, Day_2 in RUL_17, and so on. The detail methodology of the sample collection and preparation, as well as image acquisition has been described in the work [30].

3.2 Methods

3.2.1 Proposed shelf-life prediction system

The shelf life prediction of mangoes has been determined by classifying them in accordance with number of days the fruit will remain edible as particular fruit's shelf life. The detailed steps of proposed system for the prediction of shelf life of fruits are elucidated below and represented in Fig.1:

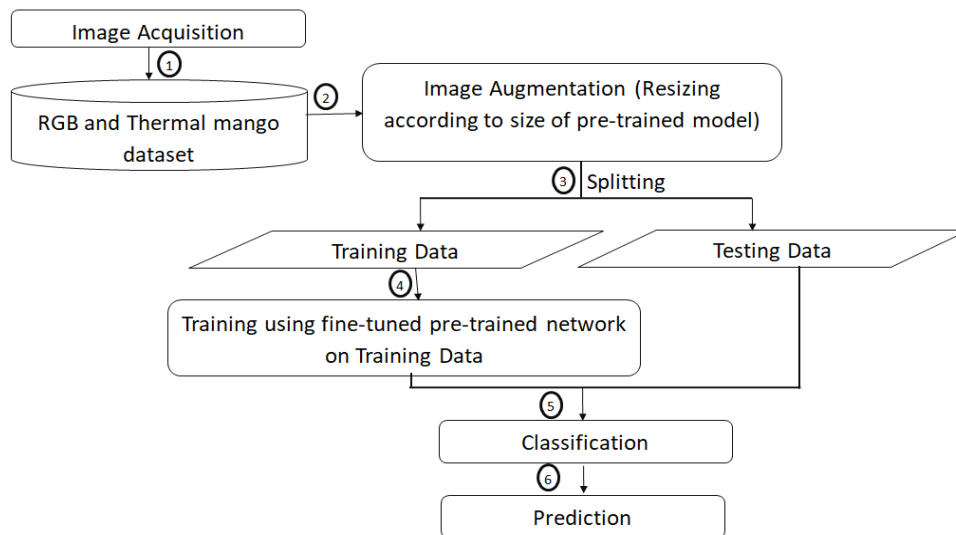


Figure 1. Proposed shelf-life prediction process.

Step 1. Image acquisition: The acquired RGB and thermal images of mangoes are divided into nineteen classes viz. RUL-1 to RUL 18 and No-Life. However, every class is not having a similar number of images. So, for the balanced distribution of images and according to the availability of images in the dataset, we have taken 240 images in every class that is total (i) 4560 images for RGB dataset and total (ii) 4560 images for thermal dataset.

Step 2. Pre-processing and data augmentation: The procured RGB and thermal images are resized according to the size of the framework used i.e. (227 x 227 x 3) or (224 x 224 x 3) which is given in Table 1. This is useful for enhancement of accuracy and speediness in training and testing phase. Moreover, to enhance the robustness of the model and to increase the generalization ability on unseen samples, normalization has been done by dividing the channels by 255 to overcome the true color of images.

Data augmentation: Although more than 10000 images are present in the dataset, this count is not sufficient. Hence, data augmentation has been employed to generate more images with no change in the labels for enlarging the dataset which results in added data [44]. The data augmentation operations consist of cropping, flipping, rotation, brightness, and contrast modification. But, this study is based on the color

resemblance of fruit, so we have not used the operations which affect the color of the image; only rotation and flipping operations have been utilized for the expansion of the training dataset by randomly augmenting the images, so the same images are never seen by the model again.

Step 3. Data Splitting: Split dataset with 75:25 proportions for training and test (i.e.validation) set of images respectively. The training set is also divided the data into batches, and performed shuffling for the rearrangement of the images randomly.

Step 4. Train the training dataset with pre-trained CNN model via transfer learning technique, where initial “weights” were used as “ImageNet” [45].

Step 5. Classify the system by applying this learned model on the test set.

Step 6. Predict the labels and its probabilities for the new set of images.

Repeat Step 2 to Step 6 for taking decision of predicting shelf life using each framework.

Table 1: Pre-trained models and their properties.

Sr. No.	Framework	Size (MB)	Depth	Parameters (Millions)	Image input size
1	SqueezeNet	4.6	18	1.24	227x227x3
2	ShuffleNet	6.3	50	1.4	224x224x3
3	MobileNetV2	13	53	3.5	224x224x3

3.2.2 CNN frameworks

In the proposed study, transfer learning-based pre-trained CNN (Convolutional Neural Network) has been adopted to forecast shelf life of mangoes for RGB and thermal imaging. The training of CNN is possible through two approaches; one is from scratch and the second is with the use of transfer learning. It is clear that the dataset used in this present work is smaller in size, the training from scratch leads to over-fitting of results on the mango shelf-life dataset used. Hence transfer learning is chosen and the use of pre-trained networks dramatically shortens the training time required for the overall computation of the results. Three CNN architectures namely SqueezeNet, MobileNetV2, and ShuffleNet have been selected for this study as they are lightweight models and elaborated in Table 1. These CNNs are pre-trained on the ImageNet dataset [45] which consists of more than 1.2 million images with 1000 image categories. This is useful for expanding the dataset which has a small number of images. The pre-trained models are able to detect generic local features such as lines, curves, or edges of a mango fruit, because these local features are almost similar in all other fruits; and these features are not specific to any class. The rest of the work is to tune the higher-level features and then these retrieved features are passed on to the classification layer, which predicts the shelf-life. Every architecture is different from each other and has its own characteristics. With the aim of better performance, speedy operations and to verify the results, these three pre-trained networks have been used for fine-tuning the parameters in this experimentation.

SqueezeNet: This model has been suggested by [46] which contains initial convolution layer continuing with eight fire modules and ended with convolution layer. The detailed architecture and parameters for SqueezeNet used in this experimentation has been adopted from [30].

MobileNetV2: The authors of [47] has proposed this model which initiates with FC (fully convolutional) layer, continued with nineteen residual bottleneck layers where depth wise separable convolution has been used. ReLU6 has been utilized in this net along with batch normalization. Also, in our entire experimentation we have employed expansion factor as 6. MobileNetV2 is the modified version of the ResNet model where 3x3 convolution is replaced with depth-wise separable convolution to create a lightweight deeper network whose size is only 14MB. The depth-wise separable convolution is a stack of 2 layers i.e. depth-wise and point-wise (1x1) convolutions. Depth-wise convolution is carried out separately for each input channel and point wise convolution is utilized for dimension reduction for solving computational bottleneck. Batch-normalization has been used in this net which accelerates the learning with very little loss.

ShuffleNet: The ShuffleNet architecture has been first put forward by [48]. It starts with convolution layer (Conv1), continuing with six shuffle unit i.e. Shuffle_Unit_2 to Shuffle_Unit_7 and at the end finished with a fully connected layer. These shuffle units which are the keystone of our fine-tuned net introduced the channel shuffle including group convolutions with the advantage of rise in representation capability and reduce in the cost of computation. The channel shuffle helps to flow the input data in the feature channels. The batch normalization has been used after each group convolution for speeding up the mini-batch training in deep networks by downsizing internal covariate shift.

So, in a SqueezeNet, a dropout with a probability of 0.5 and squeeze (1x1) and expanded (3x3) convolution has been utilized which reduces the computational cost in terms of training time. MobileNetV2 and ShuffleNet

accomplish quicker learning in comparison to SqueezeNet because of batch normalization. However, in MobileNetV2, the use of depth-wise separable convolution which divides a single 1x1 convolution into two 1x1 convolutions along with the utilization of expansion factor improves the efficiency of training time as compared to SqueezeNet. But, with an added function of channel shuffling including group convolution operation, ShuffleNet resulted in reducing the training time with greater accuracy in comparison to SqueezeNet and MobileNetV2 models. So, by considering the outperformance of the ShuffleNet amongst the three models, it has been explained in detail.

3.2.3 Proposed ShuffleNet model architecture

The architecture of proposed fine-tuned ShuffleNet model where input is all augmented images which is passed to this model is illustrated in Fig. 2 which represents the collection of Shuffle_Unit blocks clustered in three phases. The first component in every phase is tested using stride = 2 and second component including stride = 1. Then the softmax activation function has been used in the output layer for predicting the classification score of 19 classes to evaluate the shelf life. The output size has been reduced from 224x224 to 7x7 for getting high-level feature maps.

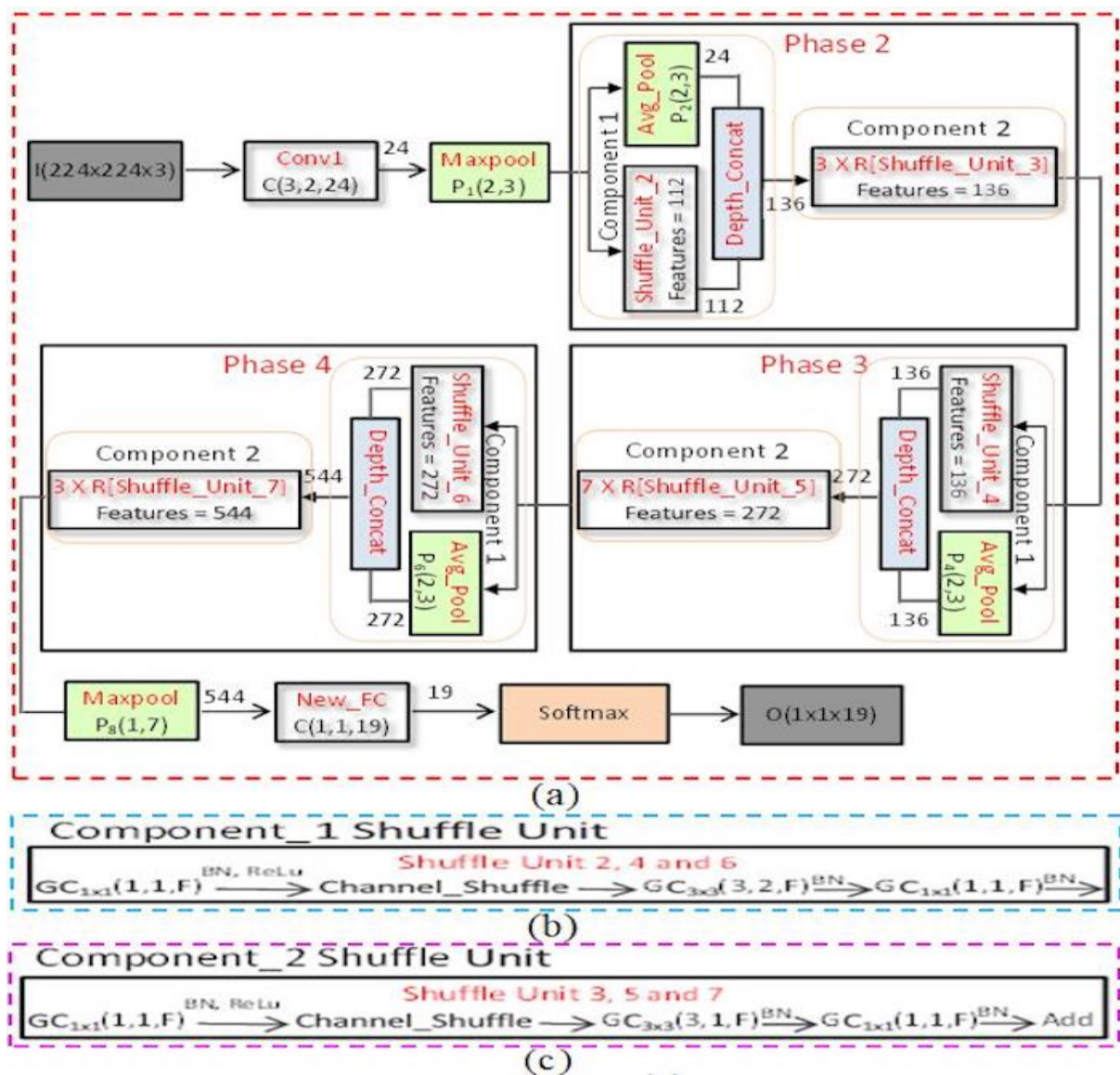


Figure 2. Comprehensive description of ShuffleNet. (a) Proposed fine-tuned ShuffleNet architecture; (b) Component_1 shuffle unit with depth concatenation; (c) Component_2 shuffle unit with elementwise addition.

Shuffle unit: In shuffle units, the numbers of features have been gradually expanded from 112 to 544 until the last unit. In component_1 and component_2 shuffle units, the first 1x1 group convolution ($GC_{1 \times 1}$) proceeds with channel shuffle then followed by 3x3 group convolution ($GC_{3 \times 3}$) again proceeds with second 1x1 group convolution ($GC_{1 \times 1}$), because of which it is possible to acquire the input data from various groups and I/P, as well as O/P, channels can be interrelated. The second 1x1 group convolution benefits to rebuild channel dimensions to meet the shortcut path of each unit. The batch normalization has been used after each group convolution and only after the first point-wise group convolution, the ReLU activation layer has been applied next, followed by batch normalization. In component_2, elementwise addition has been performed after the second 1x1 group convolution but in component_1, average pooling (3x3 mask with stride=2) has been added to the shortest route and elementwise addition has been substituted with depth-wise concatenation to extend channel dimension and inhibit possible overheads. In the current study, the number of groups (g) used in all shuffle units are 4 (i.e. $g = 4$) which regulates missing information in 1x1 group convolution and the output channels have been computed. The output channels are: {28, 34} in phase 2 for shuffle unit_2 and shuffle unit_3; {34, 68} in phase 3 for shuffle unit_4 and shuffle unit_5; {68 and 136} in phase 4 for shuffle unit_6 and shuffle unit_7 which shows the output channels have been doubled in every succeeding phase. The notations in each block of Fig. 2 are specified as follows:

- $I(h \times w \times c)$: dimensions of augmented input image as width, height and channels.
- $P(s, k)$: P is pooling layer (either max-pool or average pool) where s indicates stride and kernel size kxk.
- $C(k, s, F)$: C is convolution layer where k is the filter size, s indicates stride and F is no. of features.

In shuffle units,

- $GCKkxk$ represents group convolution with filter size kxk.
- $TxR[.]$: R signifies repeated block [.] and TxR means block repeated T times.
- $O(1 \times 1 \times 19)$: output that signifies scores of the 19 classes.

4 Results and Discussion

We perform the experiment on three CNN based pre-trained networks i.e. SqueezeNet, ShuffleNet and MobileNetV2 to forecast the shelf life of fruits for RGB and thermal imaging. The experiment has been performed using a device with 3.90 GHz CPU frequency, 8 GB RAM, and GPU with NVIDIA GeForce® GTX730M graphics card for speedy operations and run several times to finalize the proper hyper-parameters viz. batch size, number of epochs, number of iterations, validation frequency which are exemplified in Table 2.

Table 2: Empirical fine-tuned optimization.

	Batch	Epochs	Iterations	Iterations per epochs	Validation frequency
RGB Dataset					
SqueezeNet	16	20	4260	213	500
ShuffleNet	16	5	1065	213	1000
MobileNetV2	16	5	1065	213	1000
Thermal Dataset					
SqueezeNet	32	25	2650	106	500
ShuffleNet	32	10	1060	106	1000
MobileNetV2	16	5	1065	213	1000

4.1 Experimental evaluation

First, we capture and store the RGB and thermal images of mangoes for around 19 days and create the classes accordingly, like Day_1 captured images were store in class RUL_18, Day_2 in RUL_17 and so on. In all experiments, once the augmentation has been performed on both RGB and thermal datasets, they have been partitioned as training and test (i.e. validation) data, where 75% (180 images per class) in the dataset have been utilized for training and remaining 25% (60 images per class) for testing operation. So, training has been carried out with total 6840 images which contribute better effect of generalization to the network. We chose three CNN based pre-trained models for prediction of shelf life of mango fruit by comparing the outcomes with respect to classification accuracy and training time. We select to compare SqueezeNet, ShuffleNet and MobileNetV2 models, owing to their resemblance in scale i.e. all three are light weight models compare to other architectures as illustrated in Table 1. For improving results and further preventing overfitting; data augmentation, dropout, batch normalization, and L2 regularization have been utilized in the entire process. The L2 regularization ensures to

decrease the weights i.e. larger weights are penalized by preferring smaller ones. We have employed dropout with 0.5 in SqueezeNet and in MobileNetV2 and ShuffleNet, we have used batch normalization. Because of the batch normalization, ShuffleNet and MobileNetV2 perform faster learning with less loss compared to SqueezeNet. But amongst the ShuffleNet and MobileNetV2, ShuffleNet requires less computational time because of smaller model size and more accuracy by virtue of the additional functionality of channel shuffle with group convolution operation. We carry out this comparative study on our own created two image datasets: RGB and Thermal.

4.2 Optimality criterion

An optimality criterion offers insight into how good a model is by quantifying the satisfaction levels used to assess the model. Before discussing the experimental outcomes, we first represent the different fine-tune optimization parameters such as mini-batch size, epochs, iterations, number of iterations per epoch and validation frequency used in the RGB and thermal datasets for the abovementioned three frameworks in Table 2. Additionally, some common empirical parameters used in both the datasets for all three frameworks are: optimizer = adam [49], cost function = categorical cross-entropy, learning rate = 10^{-4} , first momentum = 0.9 and second momentum = 0.999 and depicted in Table 3. The proposed frameworks have been trained by Adam optimizer through categorical cross-entropy loss function for both RGB and thermal datasets. Adam is the adaptive moment estimation (whose default learning rate = 10^{-3}) which evaluates gradients first (mean) and second (uncentered variance) moments to adjust the learning rate of every weight of the network on existing mini-batch. The default values of the hyper-parameter first and second momentum have been chosen as recommended by inventors of Adam algorithm and gives better results in this study.

Table 3: Common fine-tuned hyper-parameters used in the experimentation.

Parameter	Optimizer	Loss function	Learning rate	First momentum	Second momentum
Value	Adam	categorical cross-entropy	10^{-4}	0.9	0.9999

4.3 Results

At the time of experimentation, all three models have been trained by SGD, Adam and RMSProp optimizer experimented with various learning rates. After some experiments, we could observe that Adam optimizer with a learning rate of 10^{-4} was the best configuration for the scenario. As the optimum learning rate mainly depends on the dataset we used and here we are using a transfer learning approach for fine tuning the pre-trained networks. So, the network weights need to be changed less aggressively; for which we have to select the learning rate lesser than the default one. We evaluate the results for RGB and thermal datasets based on SqueezeNet, ShuffleNet, and MobileNetV2 fine-tuned pre-trained networks. We have trained the same models 10 times ($k=10$) and the accuracy as well as training time along with deviation have been presented in Table 4. As illustrated in Table 4, average accuracy is $96.06 \pm 0.94\%$ for SqueezeNet, $96.77 \pm 0.64\%$ for MobileNetV2, and $98.01 \pm 0.44\%$ for ShuffleNet of thermal image dataset as compared with the average accuracy of RGB dataset i.e. $95.01 \pm 0.66\%$, $96.1 \pm 0.84\%$, and $97.11 \pm 0.6\%$, respectively. From the results, we can conclude that for the ShuffleNet model, the average accuracy has risen by almost 1% as compared to the MobileNetV2 model, wherein it has been raised by nearly 2% in comparison to the SqueezeNet model.

Table 4: Mango shelf life evaluation results for RGB and Thermal datasets.

Sr. No.	Dataset_CNNframework	Accuracy (%)	Time (minutes)	Speed up
1	RGB_SqueezeNet	95.01 ± 0.66	236.52 ± 36.18	
2	Thermal_SqueezeNet	96.06 ± 0.94	52.54 ± 4.05	4.5x
3	RGB_MobileNetV2	96.1 ± 0.84	110.23 ± 11.1	
4	Thermal_MobileNetV2	96.77 ± 0.64	45.14 ± 4.39	2.4x
5	RGB_ShuffleNet	97.11 ± 0.6	85.55 ± 4.16	
6	Thermal_ShuffleNet	98.15 ± 0.44	41.08 ± 2.32	2.2x

4.3.1 Results based on RGB datasets

The results explored in this are related to RGB dataset. The training profile indicates the performance results, has been portrayed in Fig. 3, Fig. 4 and Fig. 5 for SqueezeNet, MobileNetV2, and ShuffleNet respectively. It identifies the time and accuracy of SqueezeNet, MobileNetV2, and ShuffleNet. The time taken by SqueezeNet

and MobileNetV2 is larger than ShuffleNet and also the ShuffleNet shows good recognition accuracy compared to the other two nets. It also shows intermediary results of training and validation accuracy, as well as the loss over each epoch and from the Fig. 3, Fig. 4 and Fig. 5, we can conclude that the network has been trained appropriately as the loss decreases and accuracy increases consistently from start to end. Here, average loss of training iteration validation dataset decreases to 0.1727.

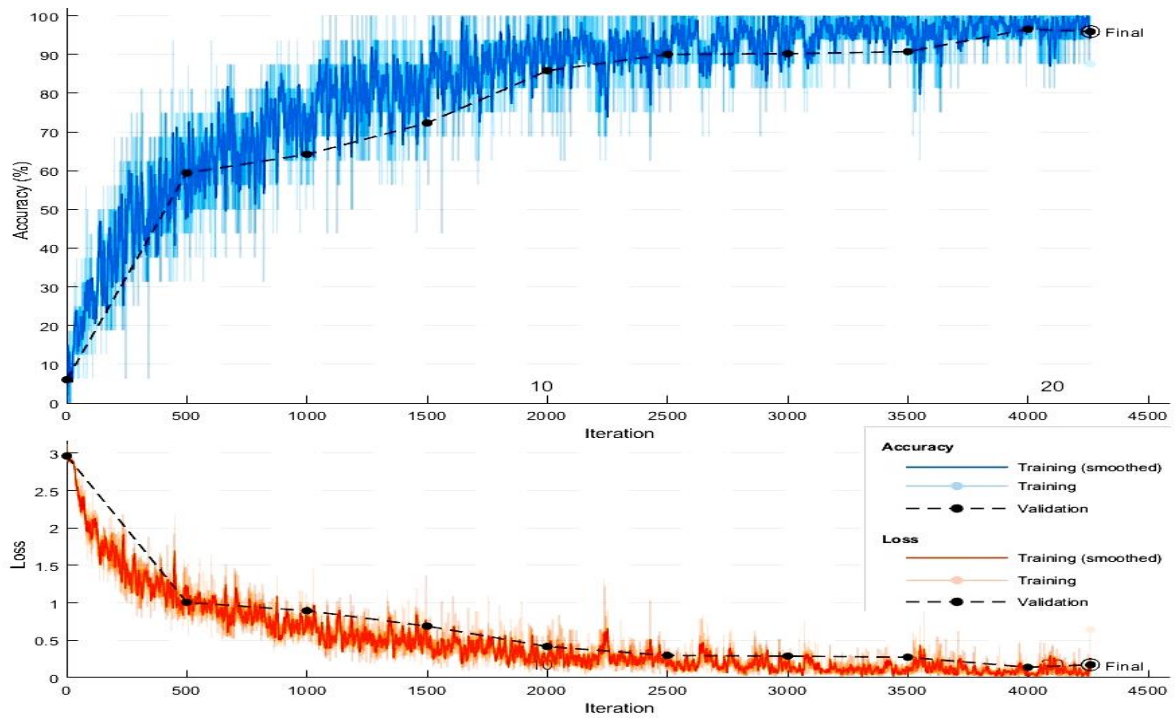


Figure 3. Training profile of SqueezeNet framework over RGB data.

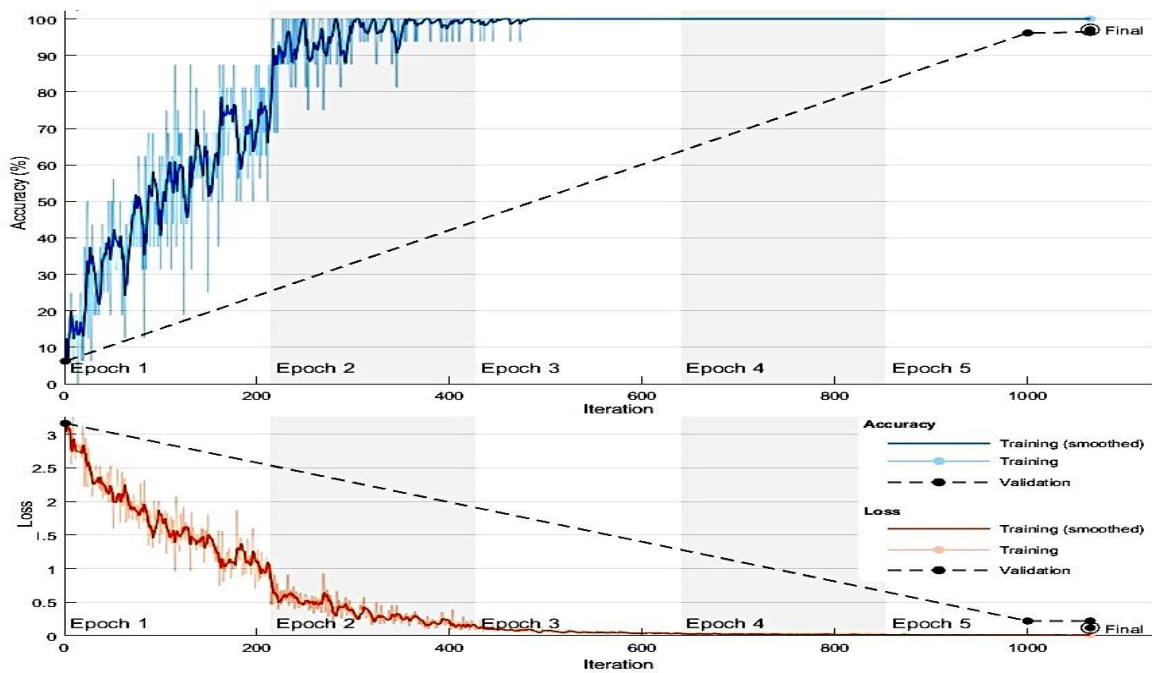


Figure 4. Training profile of MobileNetV2 framework over RGB data.

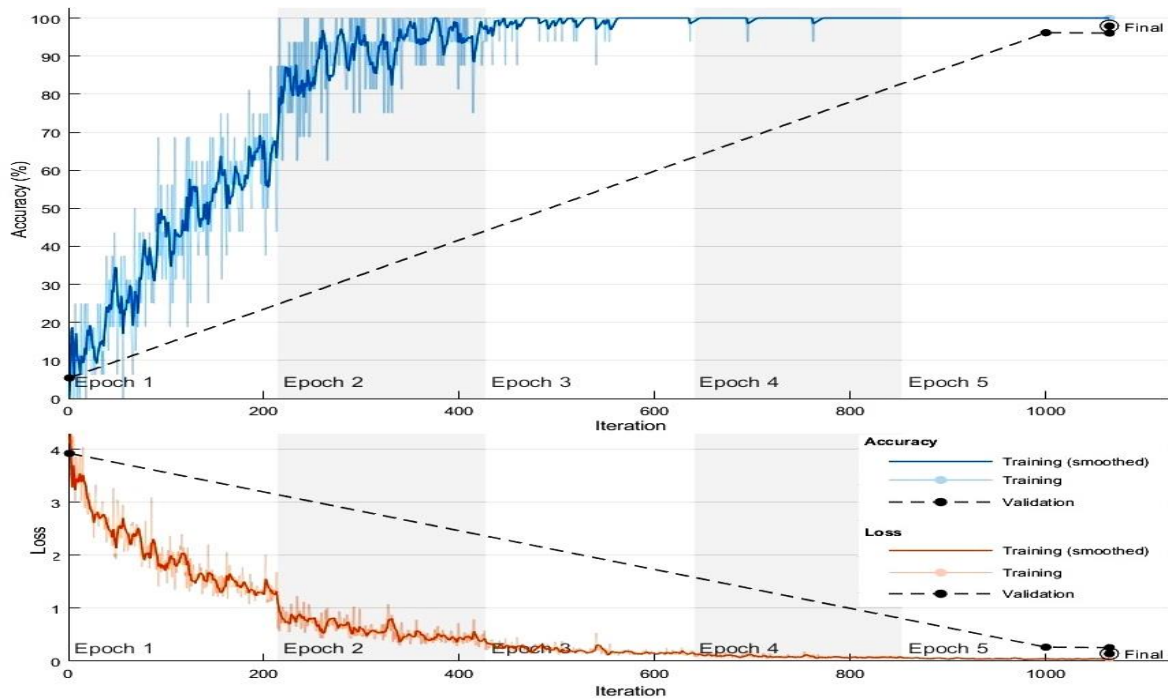


Figure 5. Training profile of ShuffleNet framework over RGB data.

The confusion matrix which is outlined in Fig. 6, Fig. 7, and Fig. 8 has been utilized as a metric function to evaluate the performance of the SqueezeNet, MobileNetV2, and ShuffleNet models over the test image dataset respectively where row epitomizes the true class and the column signifies the predicted class. The row summary represents percentage of correctly identified and mis-classified samples for every true class which are referred to as Recall and False Positive Rate (FPR) respectively. The column summary represents the proportion of correctly and mis-classified samples for each predicted class which are referred as Precision and False Discovery Rate (FDR) respectively.

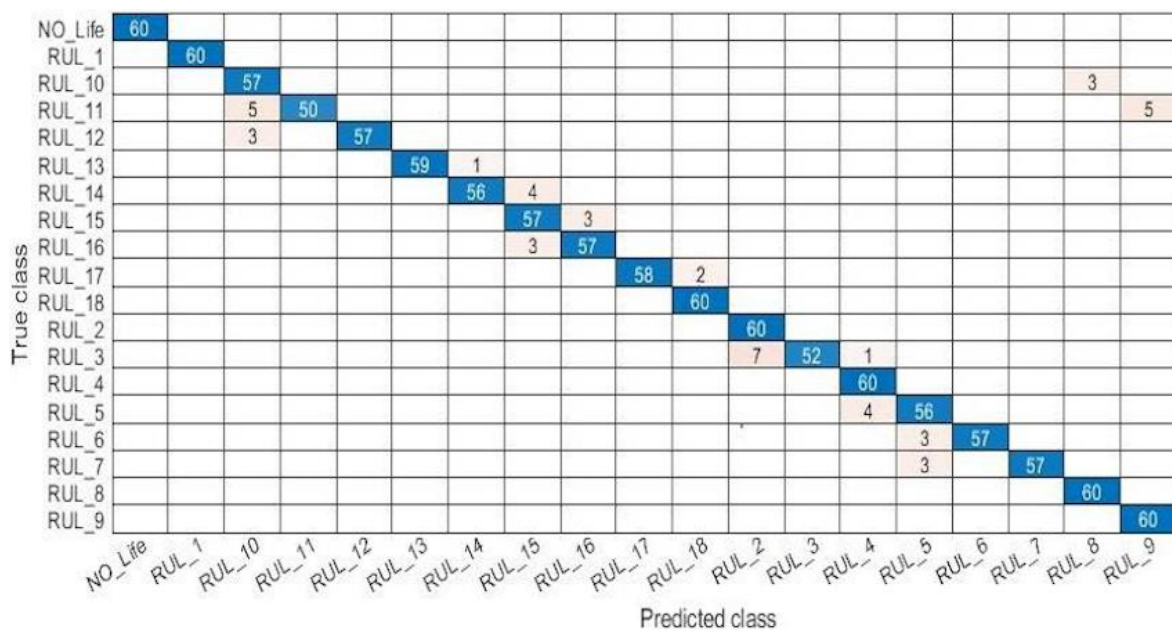


Figure 6. Confusion matrix over SqueezeNet_RGB dataset.

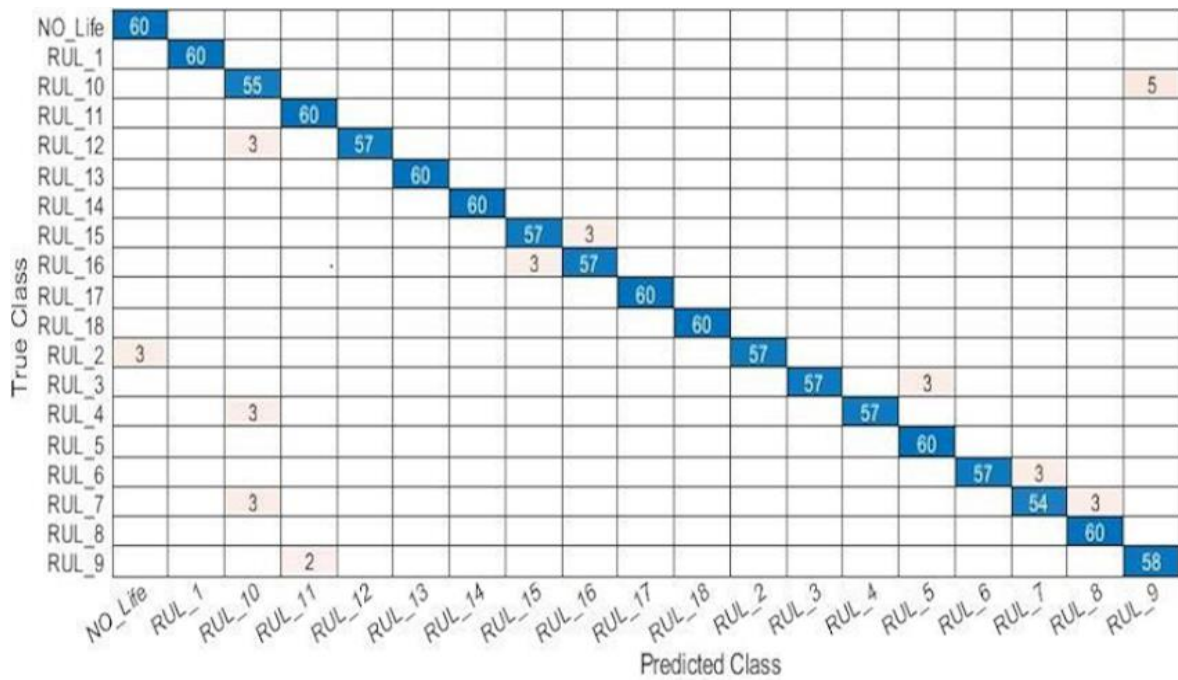


Figure 7. Confusion matrix over MobileNetV2 RGB dataset.

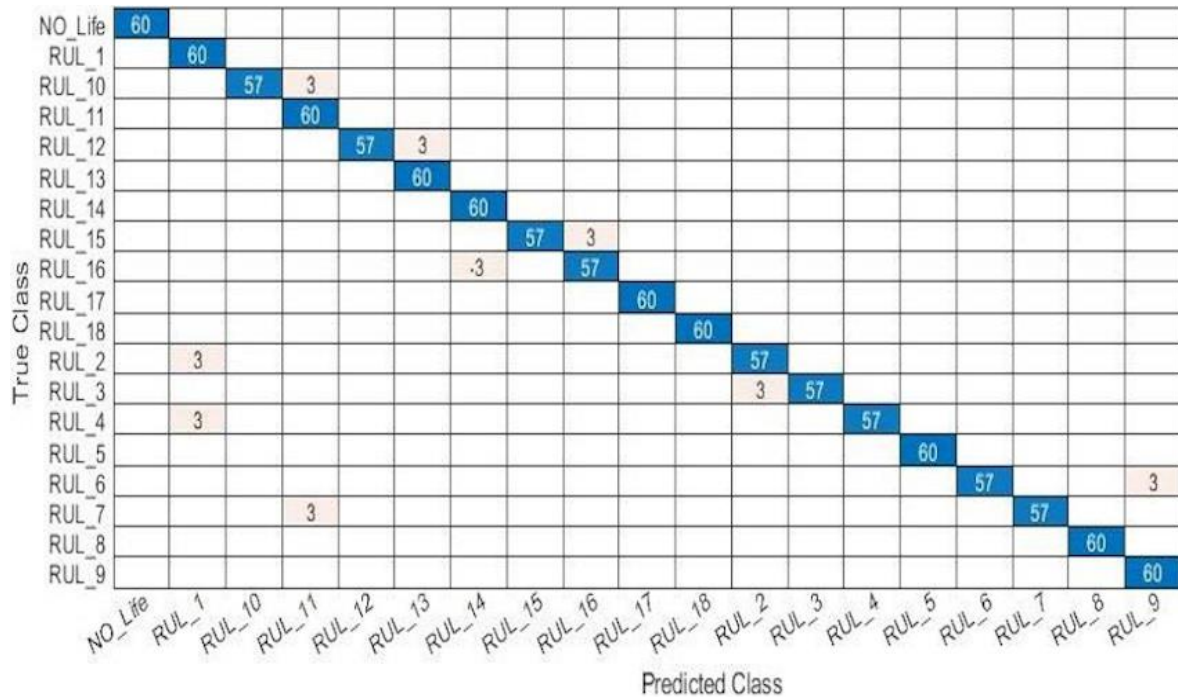


Figure 8. Confusion matrix over ShuffleNet_RGB dataset.

4.3.2 Results based on Thermal datasets

This section proposes all the results related to thermal dataset. The training profile of SqueezeNet, MobileNetV2, and ShuffleNet has been depicted in Fig. 9, Fig. 10 and Fig. 11 respectively which evaluates the performance metrics as accuracy as well as loss, number of iterations to reach maximum validation accuracy and training time. The average loss of training iteration validation dataset decreases to 0.1312.

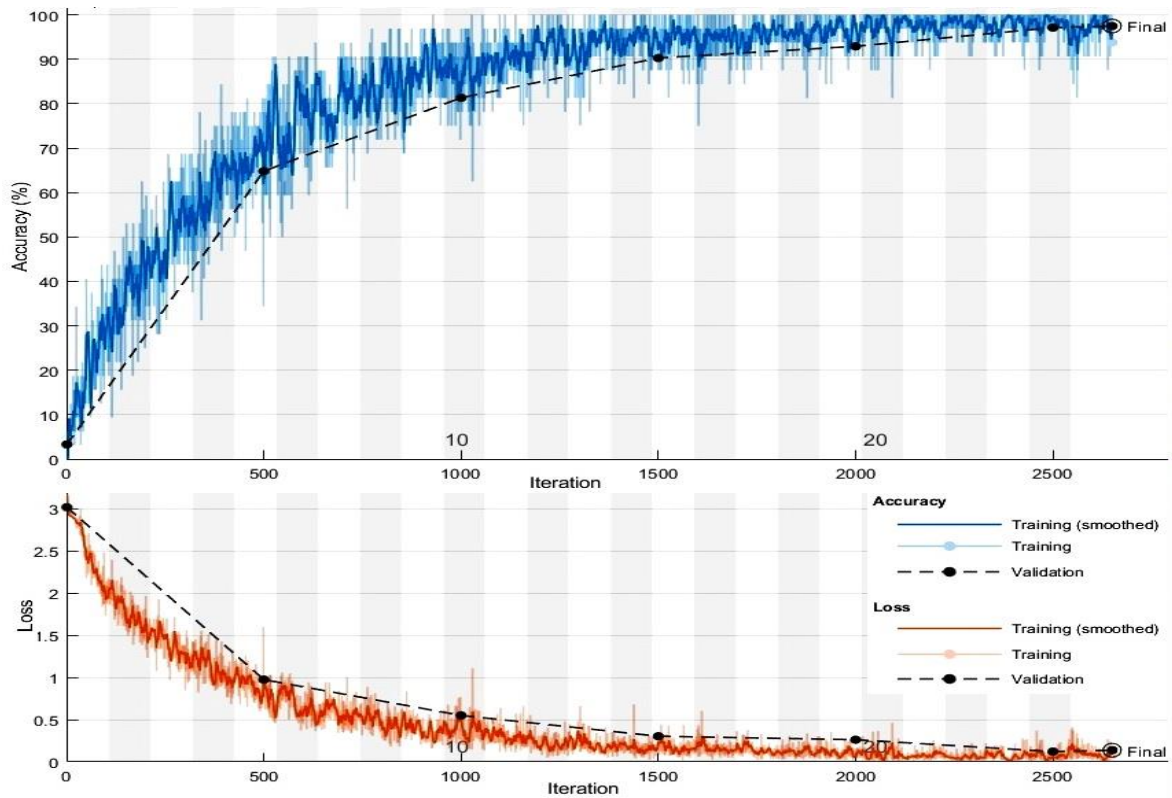


Figure 9. Training profile of SqueezeNet framework over Thermal data.

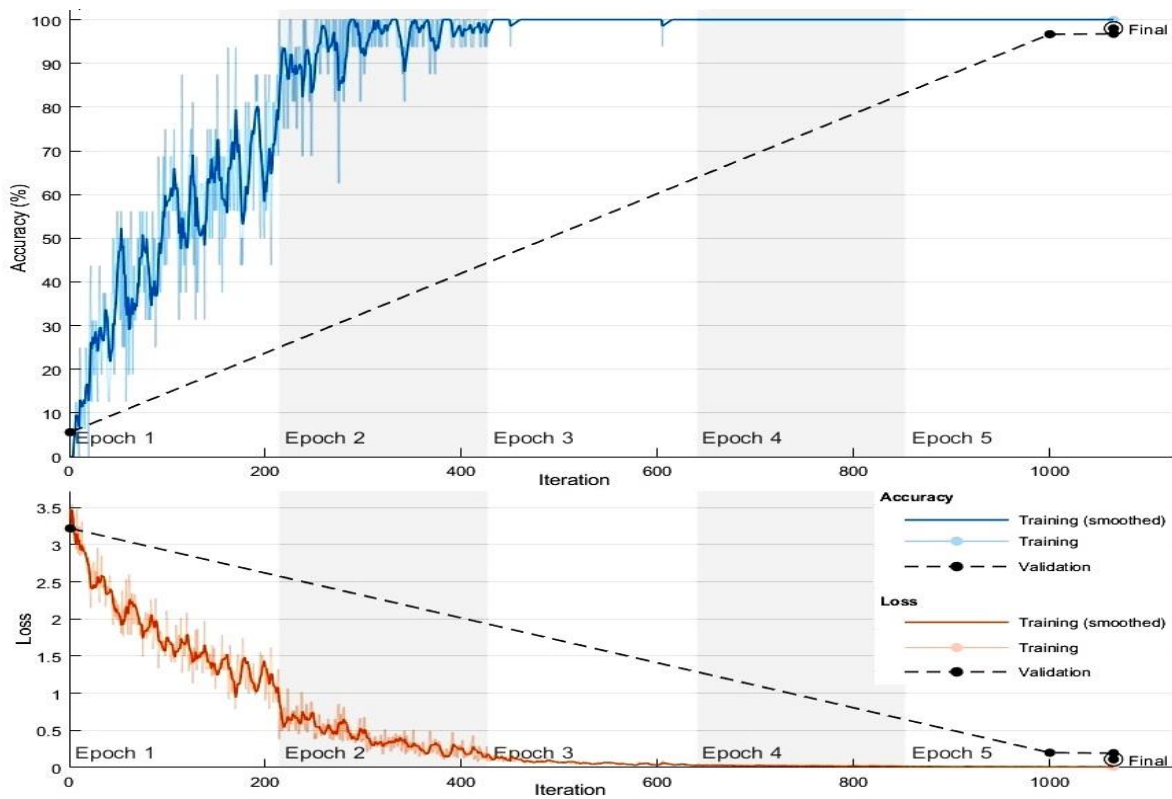


Figure 10. Training profile of MobileNetV2 framework over Thermal data.

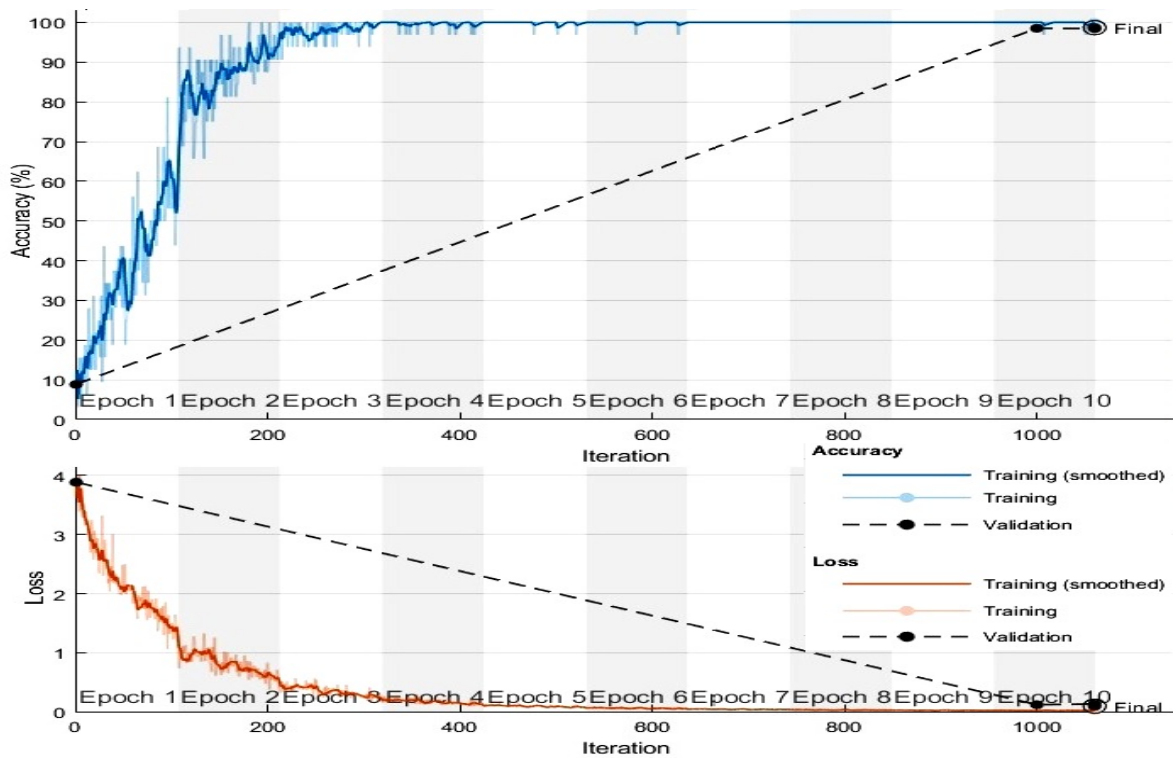


Figure 11. Training profile of ShuffleNet framework over Thermal data.

For measuring the performance level, we use some sort of statistical tool. So, we evaluate some performance measures like F1-score, precision, recall, FPR, and FDR based on confusion matrix illustrated in Fig. 12, Fig. 13, and Fig. 14 and demonstrated in Table 5. The Fig. 12, Fig. 13 and Fig. 14 depicts the confusion matrix that determines the efficacy of the SqueezeNet, MobileNetV2, and ShuffleNet across the test dataset where diagonal epitomizes correct results and from the figures, it can be observe that most of results predicted correctly.

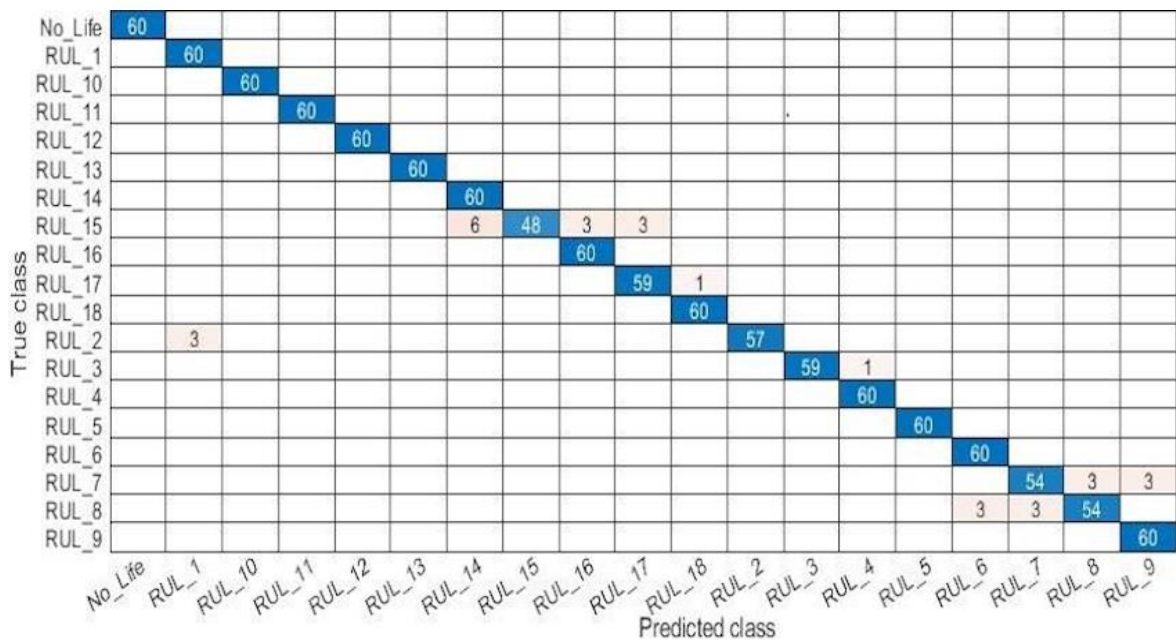


Figure 12. Confusion matrix over SqueezeNet_Thermal dataset.

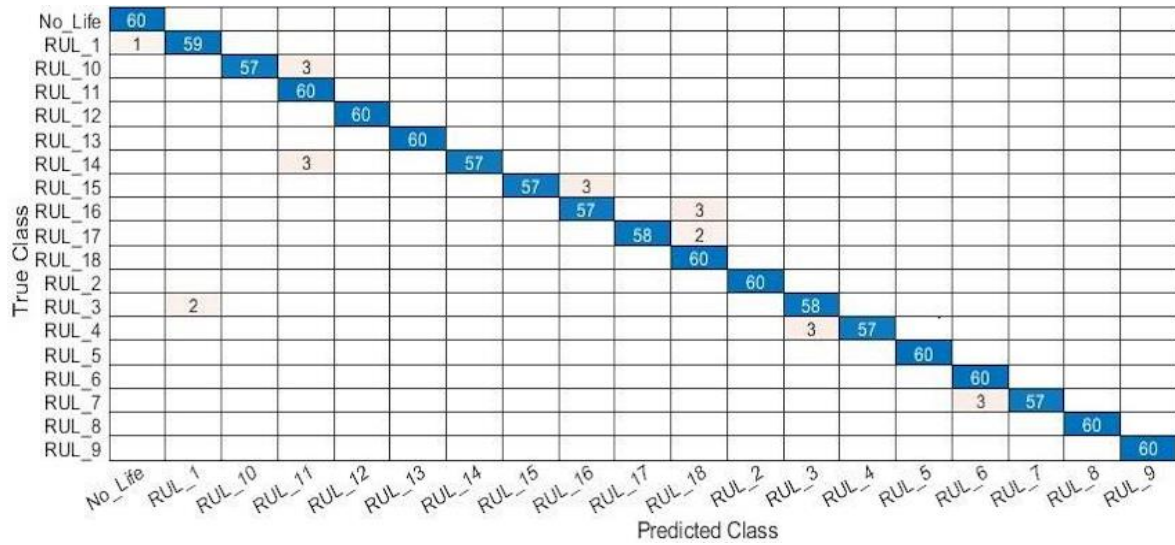


Figure 13. Confusion matrix over MobileNetV2 Thermal dataset.

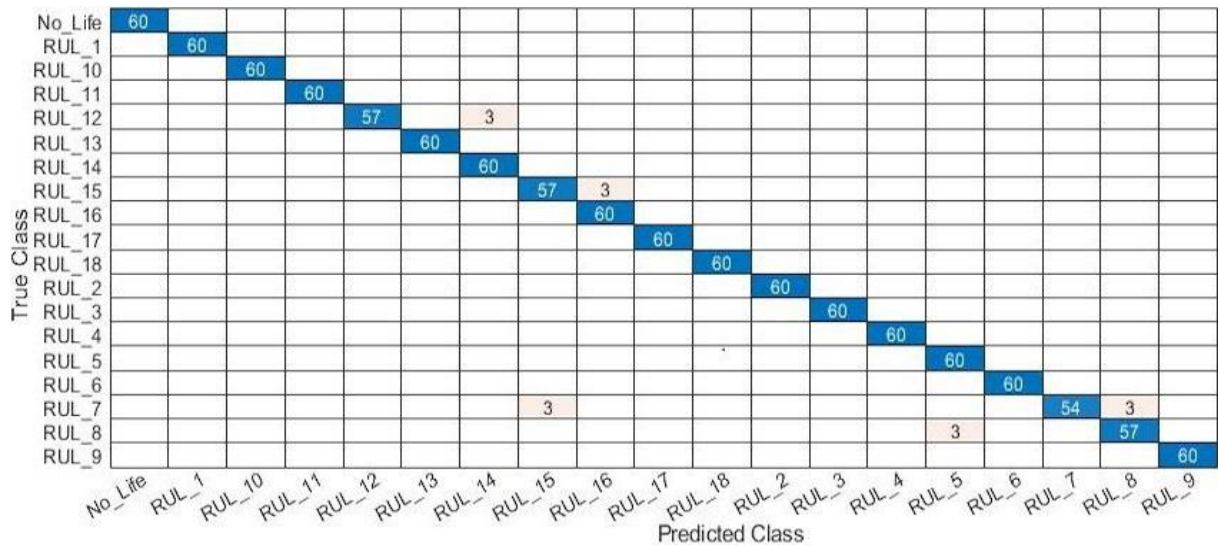


Figure 14. Confusion matrix over ShuffleNet_Thermal dataset.

Table 5: Evaluation metrics for RGB and Thermal datasets.

Sr. No.	Dataset_model	Macro AVG (%)				
		Precision	Recall	F1-Score	FPR	FDR
1	RGB_SqueezeNet	95.1	95.0	95.1	6.5	7.4
2	Thermal_SqueezeNet	96.2	95.8	96.0	8.1	4.7
3	RGB_MobileNetV2	96.1	96.1	96.1	5.7	6.1
4	Thermal_MobileNetV2	97.1	96.9	97.0	4.3	5.2
5	RGB_ShuffleNet	97.5	97.0	97.3	5.0	6.3
6	Thermal_ShuffleNet	98.3	98.1	98.2	6.3	4.9

4.3.3 Prediction/test results

In the end, to ascertain the desired shelf life of fruit by highest probability approach, the prediction has been performed on the new set of images which have not been used while training wherein the threshold has been set to

0.75 and the four sample images according to their predicted labels and probability of carrying those labels have been presented in Fig. 15 for RGB and thermal dataset.

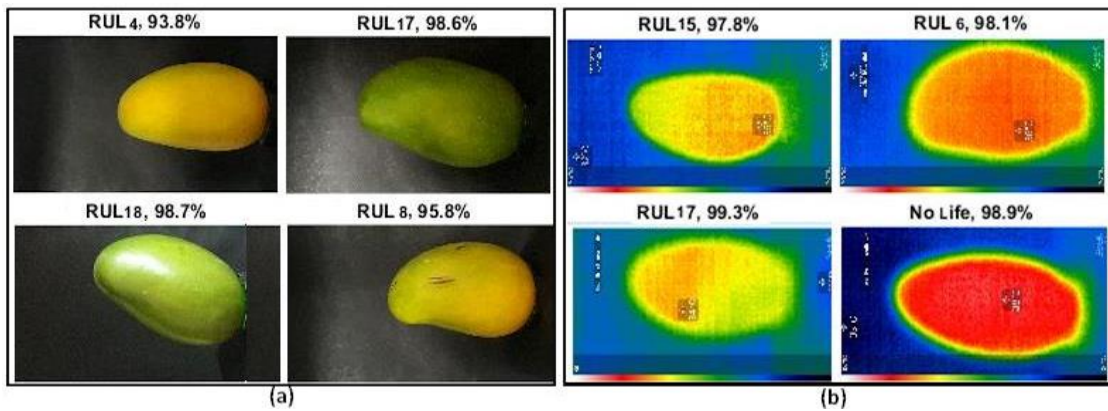


Figure 15. Predicted output of shelf-life using (a) RGB data; (b) Thermal data.

As presented in Fig. 16, the classification accuracy is $96.06 \pm 0.94\%$ for SqueezeNet, $96.77 \pm 0.64\%$ for MobileNetV2 and $98.15 \pm 0.44\%$ for ShuffleNet of thermal image dataset as compared with the accuracy of RGB dataset i.e. $95.01 \pm 0.66\%$, $96.1 \pm 0.84\%$ and $97.11 \pm 0.6\%$ respectively. It has been inferred that by implementing pre-trained networks on the RGB and thermal datasets, the most attainable accuracy is $97.11 \pm 0.6\%$ and $98.15 \pm 0.44\%$ respectively using the ShuffleNet framework that states to a better generalization performance. So, the final results are significant with thermal imaging technique with respect to RGB because the results with thermal imaging technique are based on the internal features of the fruits. But, if the thermal imaging device is not available then one can go with the proposed RGB models for the prediction of the shelf-life of the fruits.

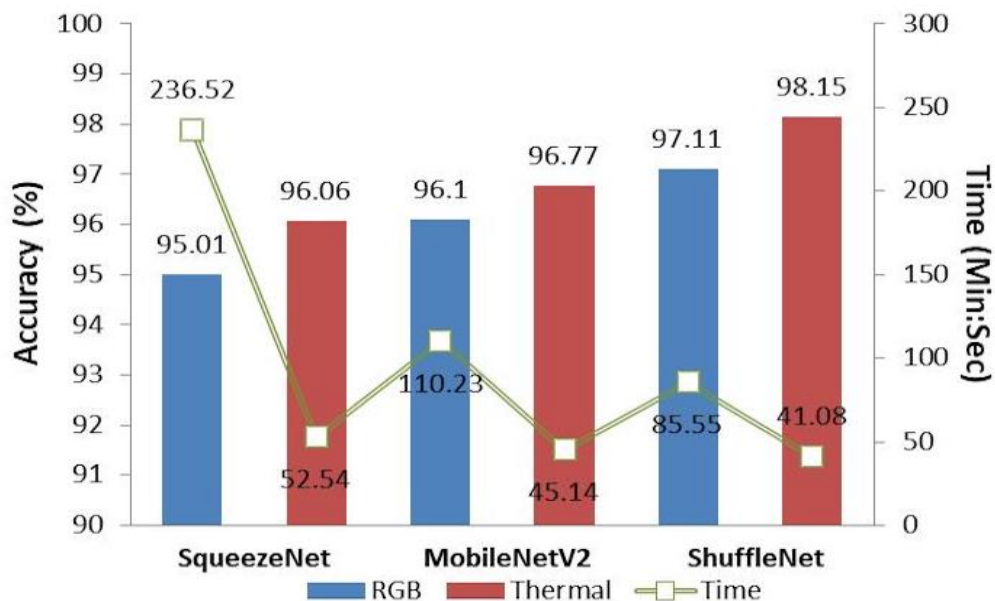


Figure 16. Evaluation metrics as accuracy and training time for RGB and Thermal data.

4.4 Comparative study

We have compared findings of present work with the other researchers [42], where experiment has been conducted on RGB images of 46 bananas to predict maturity states and shelf life. The three variants of color features have been extracted and classification has been done using seven different machine learning algorithms. The highest accuracy gained by the classifiers was around 52% which is lesser than the results obtained in the present work with RGB_ShuffleNet (97.11%) and Thermal_ShuffleNet (98.15%) and graphically expressed in

Fig. 17. In addition, we have also implemented the method (color features + machine learning classifiers) discussed in [42] and achieved an average accuracy of 90.73% and 92.47% for RGB and thermal datasets. The achieved results have been compared and shown in Fig. 17.

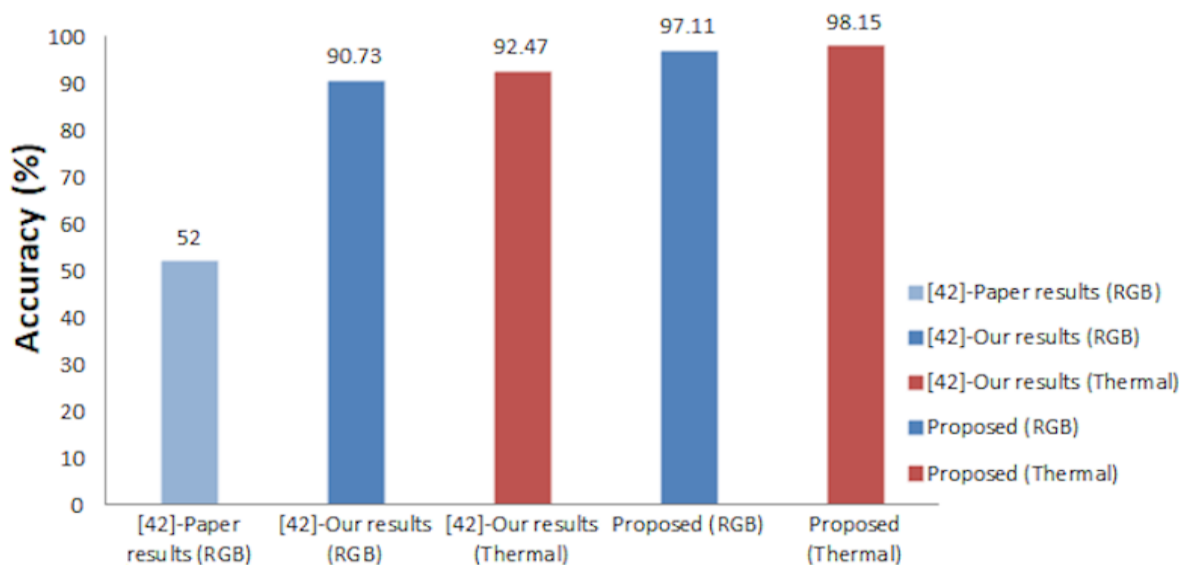


Figure 17. Comparative analysis of current work with others in the literature.

5 Conclusion

This work proposes the thermal imaging coupled with transfer learning method for evaluating the shelf life of fruits where the thermal imaging is used to reflect the intrinsic information of fruits. By applying pre-trained networks via transfer learning (SqueezeNet, MobileNetV2 and ShuffleNet) with the RGB and thermal image datasets, the thermal imaging mechanism coupled with ShuffleNet framework attains higher accuracy against RGB images as demonstrated in Fig. 16. The result signifies that in both RGB and thermal imaging, ShuffleNet outstrips the rest two frameworks with topmost accuracy (RGB=97.11±0.6% and Thermal=98.15±0.44%) including the fastest training time (RGB=85.55±4.16minutes and Thermal=41.08±2.32minutes) whereas both SqueezeNet and MobileNetV2 perform nearly the same for accuracy in both the imaging but deprived for training time in RGB imaging. So, for real-time applications, because we expect the results quickly as well as more accurately, the ShuffleNet framework could be a good choice and with this, an average accuracy has been improved by almost 2%, as compared to the other models. This verifies the aim of the proposed study, that thermal imaging can be intended to predict the shelf life of the fruits. Considering this, the proposed system is a versatile tool that could predict the shelf life with the use of both RGB and thermal imaging techniques.

By further enhancing datasets and hardware resources, there are striking opportunities for improving the accuracy along with the speed as well as contribute greater effect of generalization to the network and hence obtain better reliability. From a futuristic prospective, we can expand the current study by using diverse varieties of fruits in the dataset.

References

- [1] FAO. 2017. "The future of food and agriculture trends and challenges", <http://www.fao.org>. [Accessed: May 10, 2020].
- [2] D. Gunders, "Wasted: how America is losing up to 40 percent of its food from farm to fork to landfill", Issue Paper IP:12-06B, *Natural Resources Defence Council (NRDC)*, New York, NY, USA, 2012.
- [3] G. Kouwenhoven, V.R. Vijayender, T. Lossonczy, "Creating sustainable businesses by reducing food waste: a value chain framework for eliminating inefficiencies", *International Food and Agribusiness Management Review*, 15(3): 119-137, 2012.
- [4] FAO, *Global Food Losses and Food Waste: Extent, Causes and Prevention*, Food and Agriculture Organization of the United Nations, Rome, Italy, 2011.

- [5] K. Patel, "Studies on Destructive and Non-destructive Quality Evaluations of Mango Fruits", Ph.D. Thesis, Aligarh Muslim University, Aligarh, 2014.
- [6] K. Peleg, "Comparison of Non-destructive and Destructive Measurement of Apple Firmness", *Journal of Agricultural Engineering Research*, 55(3):227-238, 1993. doi: <https://doi.org/10.1006/jaer.1993.1046>.
- [7] V. Bhole, A. Kumar, D. Bhatnagar, "A texture-based analysis and classification of fruits using digital and thermal images", In *ICT Analysis and Applications: Lecture Notes in Networks and Systems*, Springer, Singapore, 93:333-343, 2020. doi: https://doi.org/10.1007/978-981-15-0630-7_33.
- [8] V. Bhole, A. Kumar, D. Bhatnagar, "Fusion of color-texture features based classification of fruits using digital and thermal images: A step towards improvement", *Grenze International Journal of Engineering and Technology*, 6(1):133-141, 2020.
- [9] L. Zou, S. Ming, D. Zhang, "A New Method for Rapid Detection of the Volume and Quality of Watermelon Based on Processing of X-Ray Images", In *Computer and Computing Technologies in Agriculture*, Springer, Cham, 452:731-738, 2015. doi: https://doi.org/10.1007/978-3-319-19620-6_79.
- [10] Agricultural and Processed Food Products Export Development Authority (APEDA), "Department of Commerce and Industry, Union Budget 2018-19". <https://www.ibef.org/industry/agriculture-india.aspx>.
- [11] C.A. Gunawardena, L.J. Clark, T.J. Dennis, "A spot-type defect detection and colour identification system for agricultural produce", In *Proc. of Int. Conf. on Industrial Electronics, Control and Instrumentation*, 3:2531-2534, 1991. doi: <https://doi.org/10.1109/IECON.1991.238950>.
- [12] G. Capizzi, G. Sciuto, C. Napoli, E. Tramontana, M. Wozniak, "Novel neural networks-Based texture image processing algorithm for orange defects classification", *International Journal of Computer Science and Applications*, 13(2):45-60, 2016.
- [13] K. Kheiralipour, A. Tabatabaefar, H. Mobli, S. Rafiee, M. Sharifi, A. Jafari, A. Rajabipour, "Some physical and hydrodynamic properties of two varieties of apple (*Malus domestica* Borkh L.)", *International Agrophysics*, 22:225-229, 2008.
- [14] N. Kondo, A. Usman, M. Monta, H. Murase, "Machine vision based quality evaluation of Iyokan orange fruit using neural networks", *Computers and Electronics in Agriculture*, 29:135-147, 2000.
- [15] S. Naik, B. Patel, "CIELab based color feature extraction for maturity level grading of mango (*Mangifera Indica* L.)", *Journal of System and Information Technology*, 24-33, 2014.
- [16] C.S. Nandi, B. Tudu, C. Koley, "Machine vision based techniques for automatic mango fruit sorting and grading based on maturity level and size", In *Sensing Technology: Current Status and Future Trends II, Smart Sensors, Measurement and Instrumentation*, Springer, Switzerland, 8:27-46, 2014.
- [17] C.S. Nandi, B. Tudu, C. Koley, "Machine vision based automatic fruit grading system using fuzzy algorithm", *Proc. of Conf. on Control, Instrumentation, Energy and Communication*, IEEE, 26-30, 2014.
- [18] I. Ozturk, S. Ercisli, M. Kara, Y. Erturk, F. Kalkan, "The genotypic effect on physical properties of three matured apple cultivars", *Bulgarian Journal of Agricultural Science*, 17(3): 333-338, 2011.
- [19] R. Pandey, N. Gamit, S. Naik, "A novel non-destructive grading method for mango (*Mangifera Indica* L.) using fuzzy expert system", *Proc. of Conf. on Advances in Computing, Communications and Informatics*, IEEE, 1087-1094, 2014. doi: <https://doi.org/10.1109/ICACCI.2014.6968366>.
- [20] D. Sahu, R. Potdar, "Defect identification and maturity detection of mango fruits using image analysis", *American Journal of Artificial Intelligence*, 1(1): 5-14, 2017. doi: [10.11648/j.ajai.20170101.12](https://doi.org/10.11648/j.ajai.20170101.12).
- [21] F. Sultana, S. Galib, F. Hasan, A. Jerin, "Digital measurement of maturity indices of mangoes using selected image features", *Journal of Food Processing and Technology*, 8(11):1-7, 2017.
- [22] D. SuryaPrabha, J. Satheesh Kumar, "Assessment of banana fruit maturity by image processing technique", *Journal of Food Science and Technology*, 52(3):1316-1327, 2013. doi: [10.1007/s13197-013-1188-3](https://doi.org/10.1007/s13197-013-1188-3).
- [23] O. Aiadi, M.L. Kherfi, B. Khaldi "Automatic Date Fruit Recognition Using Outlier Detection Techniques and Gaussian Mixture Models", *Electronic Letters on Computer Vision and Image Analysis*, 18(1): 51-75, 2019. doi: <https://doi.org/10.5565/rev/elcvia.1041>.
-

- [24] P. Baranowski, W. Mazurek, B. Witkowska-Walczak, C. Sławiński, “Detection of early apple bruises using pulsed-phase thermography”, *Postharvest Biology and Technology*, 53:91-100, 2009.
- [25] D. Jawale, M. Deshmukh, “Real time automatic bruise detection in (Apple) fruits using thermal camera”, *Proc. of Conf. on Communication and Signal Processing*, IEEE, 1080-1085, 2017.
- [26] Z. Jianmin, Z. Qixian, L. Juanjuan, X. Dongdong, “Design of on-line detection system for apple early bruise based on thermal properties analysis”, *Proc. of 8th Conf. on Intelligent Computation Technology and Automation*, IEEE, 47-50, 2010. doi: <https://doi.org/10.1109/ICICTA.2010.568>.
- [27] J. Varith, G.M. Hyde, A. L. Baritelle, J. K. Fellman, T. Sattabongkot, “Non-contact bruise detection in apples by thermal imaging”, *Innovative Food Science and Emerging Technologies*, 4(2): 211-218, 2003.
- [28] D.M. Bulanon, T.F. Burks, V. Alchanatis, “Study on temporal variation in citrus canopy using thermal imaging for citrus fruit detection”, *Biosystems Engineering*, 101(1): 161-171, 2008.
- [29] L.Z. Jiao, W.B. Wu, W.G. Zheng, D.M. Dong, “The infrared thermal image-based monitoring process of peach decay under uncontrolled temperature conditions”, *Journal of Animal and Plant Sciences*, 3(1):202-207, 2015.
- [30] V. Bhole, A. Kumar, “Mango Quality Grading using Deep Learning Technique: Perspectives from Agriculture and Food Industry” In *Proc. of Conf. on Information Technology Education*, ACM, 180-186, 2020. doi: <https://doi.org/10.1145/3368308.3415370>.
- [31] S. Naik, B. Patel, “Thermal imaging with fuzzy classifier for maturity and size based non-destructive mango (*Mangifera Indica L.*) grading”, In *Proc. of Conf. on Emerging Trends and Innovation in ICT*, IEEE, 16-20, 2017. doi: <https://doi.org/10.1109/ETICT.2017.7977003>.
- [32] V. Bhole, A. Kumar, “System and method for identifying fruit shell life”, Indian Patent 352727.
- [33] J. Kuzy, Y. Jiang, C. Li, “Blueberry bruise detection by pulsed thermographic imaging”, *Postharvest Biology and Technology*, 136:166-177, 2018. doi: <https://doi.org/10.1016/j.postharvbio.2017.10.011>.
- [34] T. Najeeb, M. Safar, “Dates maturity status and classification using image processing”, In *Proc. of Conf. on Computing Sciences and Engineering*, IEEE, 1-6, 2018. doi: <https://doi.org/10.1109/ICCSE1.2018.8374209>.
- [35] V. Bhole, A. Kumar, “Analysis of convolutional neural network using pre-trained squeezenet model for classification of thermal fruit images”, In *CRC Press ICT for Competitive Strategies*, Boca Raton, 759-768, 2020. doi: <https://doi.org/10.1201/9781003052098-80>.
- [36] H. Haydar, G. Ibrahim, M. Mehmet, M. Bayram, “Post-harvest chemical and physical–mechanical properties of some apricot varieties cultivated in turkey”, *J.of Food Engineering*, 49:303-310, 2007.
- [37] O.J. Oyelade, P.O. Odugbenro, A.O. Abioye, N.L. Raji, “Some physical properties of african star apple (*Chrysophyllum Alibidum*) seeds”, *Journal of Food Engineering*, 67(4): 435-440, 2005.
- [38] M. Soltani, R. Alimardani, M. Omid, “Modeling the main physical properties of banana fruit based on geometrical attributes”, *International Journal of Multidisciplinary Sciences and Engineering*, 2(2):1-6, 2011.
- [39] J.C. Fan, G.M. Zhou, “A Near infrared spectroscopy qualitative analysis of apple shelf life”, *Food and Nutrition in China*, 17:47-49, 2011. doi: <https://doi.org/10.3390/molecules200813603>.
- [40] S. Zhang, J. Xue, H. Sun, J. Zhou, “Study of *malus asiatica nakai*’s firmness during different shelf lives based on visible/near-infrared spectroscopy”, *Mathematical and Computer Modelling*, 58: 1829-1836, 2013.
- [41] C.S. Nandi, B. Tudu, C. Koley, “A Machine vision-based maturity prediction system for sorting of harvested mangoes”, *IEEE Transactions on Instrumentation and Measurement*, 63(7): 1722-1730, 2014.
- [42] Nandan Thor, “Applying machine learning clustering and classification to predict banana ripeness states and shelf life”, *Int. Journal of Advanced Food Science and Technology*, 2(1):20-25, 2017.
- [43] S. Pan, Q. Yang, “A survey on transfer learning”, *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345-1359, 2010. doi: <https://doi.org/10.1109/TKDE.2009.191>.
- [44] D. Zhao, G. Yu, P. Xu, and M. Luo, “Equivalence between dropout and data augmentation: a mathematical check,” *Neural Networks*, 115:82-89, 2019. doi: <https://doi.org/10.1016/j.neunet.2019.03.013>.
-

-
- [45] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," In *Advances in Neural Information Processing Systems*, 1097-1105, 2012. <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.
- [46] F.N. Iandola et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size", arXiv preprint arXiv:1602.07360v4, 2016.
- [47] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", *IEEE/CVF Proc. of Conf. on Computer Vision and Pattern Recognition*, USA, 4510-4520, 2018. doi: <https://doi.org/10.1109/CVPR.2018.00474>.
- [48] X. Zhang, X. Zhou, M. Lin, J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices", *IEEE/CVF Proc. of Conf. on Computer Vision and Pattern Recognition*, USA, 6848-6856, 2018.
- [49] D.P. Kingma, J. Ba, "Adam: A method for stochastic optimization", arXiv preprint arXiv:1412.6980, 2014.
-