SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

SORTING OF FRESH FRUITS USING TRANSFER LEARNING

Dr Vijay Prakash Singh[1], Dr S N Talbar[2], Deepali M Bongulwar[3]

[1]*Research Guide* ,Department of Electronics and Communication Engineering, School of Engineering, Sri Satya Sai University of Technology and Medical Sciences, Sehore, MP, India

[2]*Research Co-Guide* Department of Electronics and Telecommunication Engineering, SGGSIE&T, SRTMU, Nanded (M.S), India

[3]*Research Scholar,* Department of Electronics and Communication Engineering, School of Engineering, Sri Satya Sai University of Technology and Medical Sciences, Sehore, MP, India

[1]vijaybhabha12@gmail.com, [2]sntalbar@sggs.ac.in, [3]ddeepali2006@yahoo.com

Abstract

In the farming business, inspecting the quality of the fruit is a crucial responsibility. The study offers a thorough examination of several fruit images using deep learning for freshness evaluation. Freshness is the most important fruit quality indicator, as it directly affects consumers' physical well-being and purchase intentions. Additionally, it is a significant market pricing component. Therefore, it is important to assess how fresh the fruit is. Convolutional neural networks (CNNs) have illustrated their effectiveness in a number of agricultural applications. We applied the idea of transfer learning to the evaluation of fruit quality. The idea of reusing a previously trained CNN model to address a new issue without the requirement for extensive training datasets is referred to as transfer learning. The present study aims to compare the performance of four distinct CNN architectures, AlexNet, GoogleNet, VGG-19, and ResNet-50, to find the fastest method for separating fresh from rotting fruit. The performance of CNN architectures in the fruit quality rating system is based on accuracy and processing speed. "Fruits Fresh and Rotten," a publicly accessible dataset, is used for experimentation. There are six different fruit classifications, including fresh/rotten oranges, fresh/rotten bananas, and fresh/rotten apples. 7,630 training images, 3,271 validation images, and 2,698 testing images are available. AlexNet outperforms the other three models with a validation accuracy of 99.48% and a training duration of 5 m 30 s. As a result, AlexNet excels in both time and accuracy. Additionally, various metrics like Recall, Precision, and F1 score are employed to assess the efficiency of the models. We come to the conclusion that producers can improve fruit sorting by using the AlexNet model.

Keywords:, Convolutional Neural Network, Deep Learning, Transfer Learning, Fruits, Fruits Fresh and Rotten dataset, Agriculture, Fruit quality

1. Introduction

Agriculture is the key to the sustainability of any economy. Long-term economic growth and structural change are greatly aided by agriculture. Humans regularly consume fruits and vegetables. However, before purchasing anything from the market, humans perform a sort of visual assessment. Fruits and vegetables can be assessed visually, but this is a challenging and

DOI: 10.5281/zenodo.7259996

biased procedure because it depends on qualitative assessment and the fact that these products are influenced by a number of other elements. In addition to the item's suitability for consumption, the quality of the product might affect its price in the market. Fruit of high quality will have strong selling points. To raise the fruit's selling price, the quality of the fruit must be continually maintained. Fruit quality assessment by human senses frequently results in inconsistent quality and takes a lot of time. Due to outside factors including exhaustion, resentment, and bias, visual fruit classification and quality rating are inaccurate. Despite skilled workers, there are errors in the classification of fruits in the fruit sector due to differences in visual perception. In order to examine the fruits and provide more accurate information, an automated method is needed. Automatic fruit grade classification is incredibly helpful for classifying the harvest produce according to the fruit's quality. This quality grading classification can be used to set prices and fulfil orders with specific quality criteria, as well as for other post-harvest processing. In reality, a number of characteristics, including the products' appearance, colour, texture, nutritional content, and flavour are frequently used to assess the freshness and quality of fruits and vegetables. The first three criteria have always been simple for humans to measure. It is acceptable to develop a machine learning program that can assess these elements and determine the quality of the produce. As we collect images of the fruits and categorize them into two categories, fresh or rotten for our application, these three characteristics are the ones that most significantly contribute to the quality identification. In a variety of applications, numerous researchers have employed CNN for feature extraction. The reliability of CNN feature extraction has been fairly good. However, Deep CNN shares a flaw with other Deep Learning techniques in that its training process is time-consuming and dependent on the availability of big datasets. Transfer learning has been suggested in numerous research as a solution to these issues. Transfer learning approaches are effectively used in numerous real-world settings to address NLP issues, image classification issues, name-entity identification issues, and other issues. The GoogLeNet network was employed in the study [1] to forecast banana freshness. The experimental findings demonstrated that the model successfully classified the fresh bananas, with an accuracy of 98.92%. Large image datasets were processed by the authors of [2] using well-known CNN architectures, and results were expanded to include many sugarcane types using transfer learning techniques. Results from transfer learning utilizing the GoogLeNet, AlexNet, VGG-16, and ResNet101 structures were compared to those from traditional techniques. They were looking for the quickest method for separating excellent billets from damaged ones. With AlexNet, the best results were obtained in terms of accuracy and speed. In the study [3], the researchers used a convolutional neural network to develop a system for classifying fruits and vegetables from Colombia. By considering the use of fruits and vegetables packaged in bags as a natural source of variation, this dataset varies from other fruit datasets and adds yet another level of complexity to the system. The CNN model trained using transfer learning achieves accuracy on par comparable to best in class for the Colombian fruits dataset. A deep learning-based method for fruit categorization and defect identification is suggested by the research [4]. Three datasets of varied sizes and complexity were tested using the AlexNet and MobileNetV2 models. The

DOI: 10.5281/zenodo.7259996

MobileNetV2 model outperformed the others on all three datasets, identifying the apple fruit's type and defects with 100% accuracy. The paper [5] suggested a strategy for categorizing papaya fruits according to their maturity status using machine learning and transfer learning. Seven pre-trained models were included in the transfer learning strategy. Once more, the classification model using the transfer learning strategy, when training was stopped early, VGG19 performed well, with 100% accuracy and 1 m 52 s time for training was taken into account. By applying transfer learning techniques, the authors of [6] expanded their earlier work on automatic fruit recognition based on DCNN. They created a pre-trained model with 44,406 images from 15 distinct categories. The pre-trained model, which had 15 different fruit types, was expanded using two alternative ways. As the first method's pre-trained network's first five layers, the weights were not updated when training the model for a novel class of fruit. Only convolutional layers in the second method are frozen and not fully connected layers. When they prepared the model for another class of cherry classification that was not included in their pre-trained model, they obtained a great performance of 95% identification accuracy for cherry fruit after freezing the top few layers in the pre-trained network. According to the study [7], when determining fruit ripeness, deep learning with transfer learning consistently outperforms machine learning with conventional feature extraction. Also, it demonstrates how the regularizer lessens overfitting and enhances the efficiency of the system in aspects of recall, precision, accuracy, and F-measure. According to this study, the greatest strategy for reducing overfitting in transfer learning is a dropout. The authors of the study [8] utilized the concept of pre-trained networks to address the issue of determining the quality of vegetables and fruits. Utilizing a dataset of 12 categories with fruits, vegetables, and bread, the eight pre-prepared models were assessed. Each pre-trained model underwent the process of fine-tuning in order to prepare it for training on the obtained dataset. The accuracy of these models was evaluated throughout training. Each model was trained using both the original dataset and the additional data. Based on the augmented dataset, the findings revealed that the ResNet18 showed a validation accuracy of 91.37%, whereas VGG19 showed a validation accuracy of 91.50% over the used original dataset. The suggested approach used pre-trained models like VGG19 to separate the three most prevalent diseases of grape leaf from healthy ones. When compared to manual disease identification, these models' mean accuracy on test data was 98%, demonstrating the viability of the neural network technique [9]. The research [10] evaluates convolutional neural network structures for strawberry quality examination. The six architectures evaluated were the proposed CNN, AlexNet, GoogleLeNet, VGGNet, Xception, and MobileNet. The presentation of different models was evaluated using a dataset with four class categories. They discovered from the experiment that VGGNet achieves maximum accuracy.

We have investigated and assessed CNN models like AlexNet, GoogleNet, VGG19, and ResNet50 to recognize fresh and rotten fruits. We used the publicly accessible dataset "Fruits Fresh and Rotten" to train these models. These models have been assessed based on processing speed and validation accuracy. Additionally, we have contrasted their performance based on precision, recall and F1 score. We have implemented the model using MATLAB R2022a.

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

The following sections make up the full paper: The requirement for fruit quality detection and the related work are outlined in Section 1. The information of the dataset is given in Section 2. The pre-trained CNN models and the idea of transfer learning are described in Section 3. The experimental strategy and performance evaluation are presented in Section 4. The outcomes of our experiments for each model are shown in section 5 and discussed in section 6. Section 7 gives the conclusion.

2. Dataset

The dataset "Fruits fresh and rotten" is publicly available and downloaded from WWW.Kaggle.com. This dataset was created by Sriram Reddy Kalluri [11] and updated on Aug 24, 2018. It has six fruit categories, fresh or rotten apples, fresh or rotten bananas, and fresh or rotten oranges. Instead of using the dataset as is, these images are subjected to unique augmentations. Our dataset has 10,901 images in total. The dataset was subdivided into 30% and 70% for validation and training respectively. 7,630 training images, 3,271 validation images, and 2,698 testing images are available as shown in table 1. The images have been downsized to 224X224 pixels. Data augmentation improves CNN's performance and prevents overfitting. On the training images, augmentations including scale, translation, and rotation have been made. Figure 1 shows 9 images that were chosen at random.

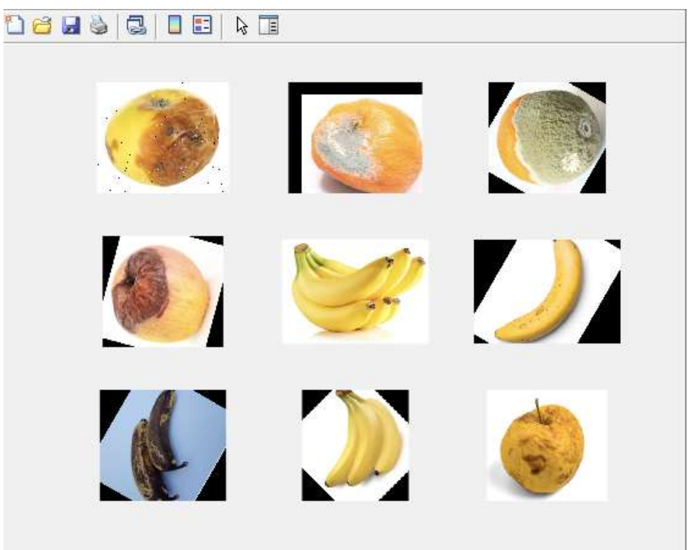

Figure 1. Sample input images

Table 1. Dataset

S. No.	Label	Count
1	Fresh Apples	1693
2	Rotten Apples	2342
3	Fresh Banana	1581
4	Rotten Banana	2224
5	Fresh Oranges	1466
6	Rotten Oranges	1595
7	Total Images	10,901
8	Training Set Size	7,630
9	Validation Set Size	3,271
10	Image Size	224X224

3. Transfer Learning

Deep learning applications frequently employ transfer learning. It is referred to as knowledge transfer. This technique is used to supplement or eliminate the necessity for the labelled dataset collection. In practical applications, it is very difficult to compile massive datasets for all feasible classes. Transferring knowledge from existing and related work facilitates the learning of a new one. Learning a new skill can be facilitated by using a pre-trained network. Comparing starting from scratch and training a network with randomly initialised weights to transfer learning, network fine-tuning is often faster and simpler. Using

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

fewer training images, we can easily apply learnt features to a new task. The four well-known CNN architectures chosen have already been pre-trained by using datasets containing millions of images divided into about a thousand classes.

We reduced the number of neurons in the final fully connected layer of four different structures AlexNet, VGG-16, GoogLeNet, and ResNet50 to six in order to more accurately represent our six classes instead of 1000. The learning could be enhanced by using those four pre-trained well-known models rather than training the models from scratch. We apply the knowledge gained by these architectures to the challenge of determining the quality of fruit. The chosen CNN architecture is retrained using the dataset. This technique for passing down knowledge is referred to as "transfer learning" or "self-taught learning." Numerous real-world applications, such as identification problems, NLP problems, picture classification problems, and others, have used transfer learning techniques.

In the current study, transfer learning was used to modify the convolutional neural network (CNN) models that had already been trained. In order to minimize the parameters and computing costs involved in the process of training, four pre-trained networks are taken into consideration for the evaluation of their performance in fruit quality detection.

3.1. **AlexNet**

Alex Krizhevsky and his colleagues proposed the AlexNet model in 2012. This paper highlights the significance of neural network depth, which has significant effects on training computation efficiency (via GPU). Training is advanced by utilizing a strong GPU execution of the convolution tasks. The suggested neural network system was placed in the top 5 among participants in the ILSVRC-2012. Many tasks, including object identification, image segmentation, and video classification, have been carried out using AlexNet. As seen in Figure 2 [12], AlexNet has a rather straightforward structure. There are eight learnable layers in the Alexnet. It comprises five layers
of convolution and three layers which are fully connected. Convolutional layers are succeeded by layers with max-pooling. A filter of size 11x11 and a stride of 4 on the first layer, 5 on the second layer, and 3 pixels on the following levels is used in the convolutional process. A softmax function receives input from the final fully connected layer's output. Each convolutional layer uses Rectified Linear Unit (ReLU) function for activation. To reduce the overfitting of image data, AlexNet employs data augmentation and dropout. Dropout is a procedure that turns each hidden neuron's output value to zero in order to simplify the co-adaptations of the neurons. With a probability of 0.5, AlexNet employs dropout in the top two fully connected layers. We employ the network to categorise a massive number of classes in accordance with our requirements. The accuracy rate is calculated once the model is trained and categorised.

DOI: 10.5281/zenodo.7259996

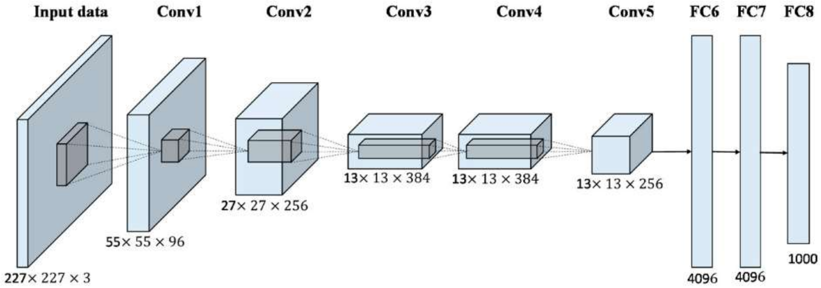

Figure 2. AlexNet Architecture [12]

3.2. **VGG19 Architecture**

The Visual Geometric Group, sometimes known as VGG, was created by Simonay and Zimmerman. It is believed to be the most widely used architecture. The VGG model has a high classification rate and a high recognition rate, making it a typical and effective CNN. It is more complex than the conventional CNN architecture from the past. In order to train the VGG19 convolutional neural network, lakhs of images from the ImageNet dataset were used. The 19-layer network can categorize images into thousand different item types. This has led to the network having complete feature representations for a variety of images. There are various configurations, VGG16 or 19, depending on the networks' convolutional layer count. VGG19 is the method we utilized in this investigation. VGG19 beats AlexNet by using multiple 3x3-sized filters in place of large-sized filters like 11x11 and 5x5. Again, it improves AlexNet by adding network depth. A fixed RGB image with dimensions of 224X224X3 serves as the conv1 layer's input. Convolutional filters are applied with a constrained 3X3 receptive field as the input image is processed through a series of layers of convolution. One of the configurations additionally employs 1X1 convolutional filters, which may be thought of as an input transformation that is first linear and then nonlinear. One pixel stride is fixed. The input convolutional layer's spatial padding is designed to maintain the spatial resolution after convolution layers, this means that the padding is 1 pixel. With stride 2, max-pooling is carried out over a 2X2 pixel window. The next step is a Rectified linear unit (ReLu), which introduces a non-linear twist to improve classification and computation time. This one performed far better than earlier models that used tanh or sigmoid functions. The reduction of volume size is handled by max-pooling. The following layer has 1000 channels for 1000 categories and is succeeded by two fully connected layers with a total of 4096 nodes each, and a softmax function as the last layer. Figure 3 shows the VGG19 architecture.

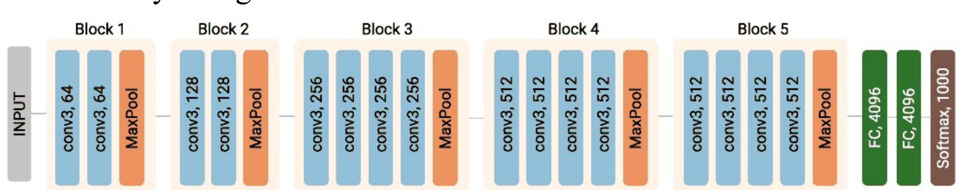

Figure 3. VGG19 Architecture [13]

3.3. **GoogleNet Architecture**

Deep convolutional neural network architecture referred to as GoogLeNet was introduced by

ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

Szegedy et al. The model, which was created using ImageNet data training is great. The architecture was chosen because it effectively identified fruits and vegetables. Figure 4 depicts the GoogleNet network's organizational structure. We retain all levels up to the output layer and transfer them to a new layer to solve the issue of a new classification. To adapt it to the new classification goal, substitute a fully connected layer, a softmax layer, and a classification output layer for the preceding layers. Scale the fully connected layer to fit the new data's number of classes in the new data. This value is set to six because it is our intended class number. In order to speed up learning in the fully connected layer relative to the transfer layer, the values of the bias learning rate factor and weight learning rate factor were increased. With the modifications, the network will be suitable for our dataset and the network training process will go more quickly. GoogleNet includes inception modules that increase both the depth and width of the networks, improving accuracy while maintaining a consistent processing load. The most widely used technique for improving the performance of deep neural networks is to expand the network size. It comprises the network's depth and width. There are certain disadvantages to the network's expanding scale. It would demand more computer resources because there would be more parameters to train. Even inside the convolutions, these issues can be resolved by switching from fully connected to sparsely connected designs. It uses the inception module to solve problems. The 1x1, 3x3, and 5x5 convolutions are used in parallel by the inception module. Prior to costly 3x3 and 5x5 convolutions, computation of reductions used 1x1 convolutions. In Figure 4 , a single inception module may be shown. When only counting layers with parameters, the GoogLeNet architecture is 22 layers deep and uses 9 inception modules. If we only count layers with parameters, there are five pooling layers in addition to the 22 deep network layers, of which there are four max-pooling and one average pooling layer. Average pooling with a 5x5 filter size and stride 3 is employed before the classifier. A dropout layer with a 70% dropout ratio is employed. All convolutional layers, even those found inside the inception modules, employ ReLU. Figure 5 shows the GooLeNet architecture's full schematic.

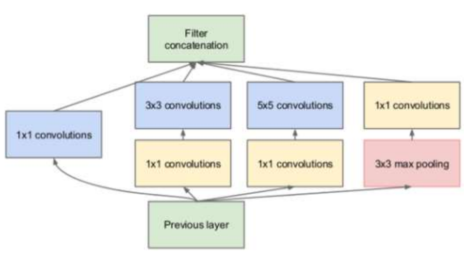

Figure 4. Inception Module [14]

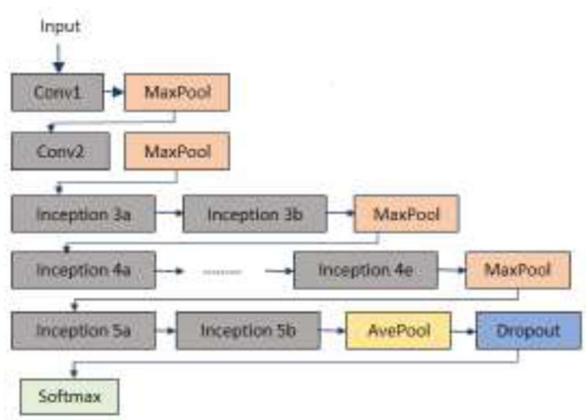

Figure 5. GoogLeNet Architecture [15]

3.4. **ResNet50 (Residual Neural Network) Architecture**

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

ResNet-50[16] is a deep residual network. The 50 denotes how many layers are present. ResNet is the most widely used model for image categorization. The skip connection is ResNet's key technological advancement. The depth of the ResNet50 Architecture is dependent on the number of consecutive modules being used rather than being fixed. The network becomes more challenging to optimize when the network's depth is increased in order to achieve more precision. ResNet fixes this issue by replacing the original application with a modified version and by introducing several connections across layers. These new connections carry out an identification operation or a 1X1 convolution while skipping multiple layers. The residual block is the fundamental building block of this network. A 1X1, 3X3, and 1X1 convolution, together with a connection that connects the input of the first convolution to the output of the third convolution, make up the network when it has 50 or more layers. ResNet variations differ in the number of layers (50, 101, or 150), but are identical in terms of architecture, pooling techniques used and filter usage.

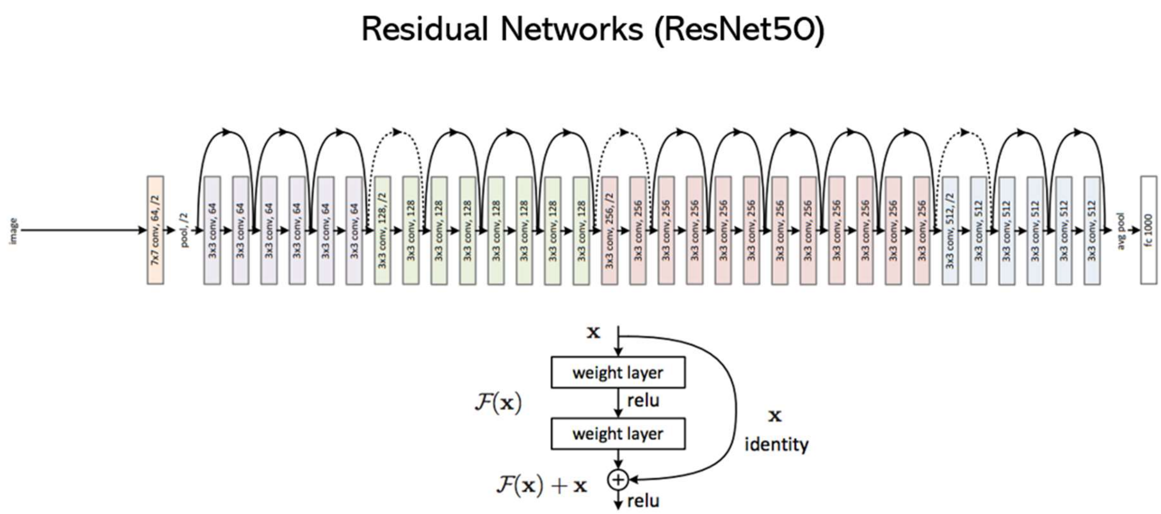

Figure 6. ResNet50 Architecture [17]

4. Materials and Methods

4.1. Experimental Setup

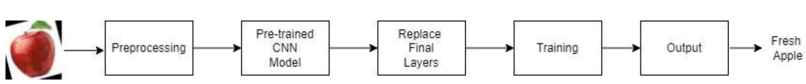

Figure 7. Block diagram

The block diagram represents the steps in detecting the quality of fruits. Initially, the images in the dataset are preprocessed to resize images to 224X24X3 pixels. We are experimenting with four CNN Models AlexNet, GoogleNet, VGG19 and ResNet50. We select one model every time and replace all the final layers. The pre-trained network's fully connected layer and classification layer are set up for 1000 classes. We swap out these two layers for new ones that are appropriate to the new dataset in order to retrain a previously trained network to categorise

DOI: 10.5281/zenodo.7259996

new images. However, we had to categorize it into six classes, thus we created a six-unit output softmax layer. The softmax layer establishes the probabilities for each class to which an input image might belong. The trained model can distinguish between fresh and rotten fruit.

4.2 Hyperparameters

Hyperparameters are variables that specify a convolutional network's architecture and regulate the learning process. Prior to training, we can decide on these values, and even after the training is complete, these values will not change. As a result, choosing the appropriate hyperparameters for a particular dataset is crucial because doing so will immediately impact the model's performance when it is used to train the model. To improve CNN performance, many hyperparameters can be changed, including learning rate, epochs, optimizer, batchsize, number of layers, and activation functions. Because it is popular and effective, we choose the stochastic gradient descent with momentum (SGDM) optimizer to classify images in CNN. The number of epochs is 6, the batchsize is 64, and the learning rate is 0.0001.

Table 2. Hyperparameters

S. No.	**Hyperparameters**	**Values**
1	Learning Rate	0.0001
2	Batchsize	64
3	Epochs	6
4	Optimizer	SGDM
5	Train-Validation Split Ratio	70%-30%
6	Bias Learn Rate Factor	20
7	Weight Learn Rate Factor	20

4.3 Experimental Framework

We utilized an Intel Core i7-8700K CPU running at 3.70 GHz and an NVIDIA GeForce RTX 2060 Super GPU with 32.0 GB of RAM. On Windows 10 Pro, we used MATLAB R2022a and the Deep Learning Toolbox to implement the model.

4.2 Performance Evaluation

The models are assessed using metrics like precision, accuracy, recall, and F1 score. The accuracy metric is utilized to assess the categorization performance of the model. It indicates the percentage of correctly categorized samples among all samples examined and calculated as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+F} \tag{1}$$

where TP denotes true positives, those who belonged to the class and were appropriately classified in the class; TN for true negatives, those who did not belong to the class but were appropriately classified in another class; FP for false positives, those who did not belong to the class but were incorrectly assigned to the class; and finally, FN for false negatives, those who belonged to the class but were incorrectly classified in another class. The number of predictions

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

made by a model for each class, as well as the classes to which those predictions belong, are summarized in a confusion matrix.

5 Experimental Results

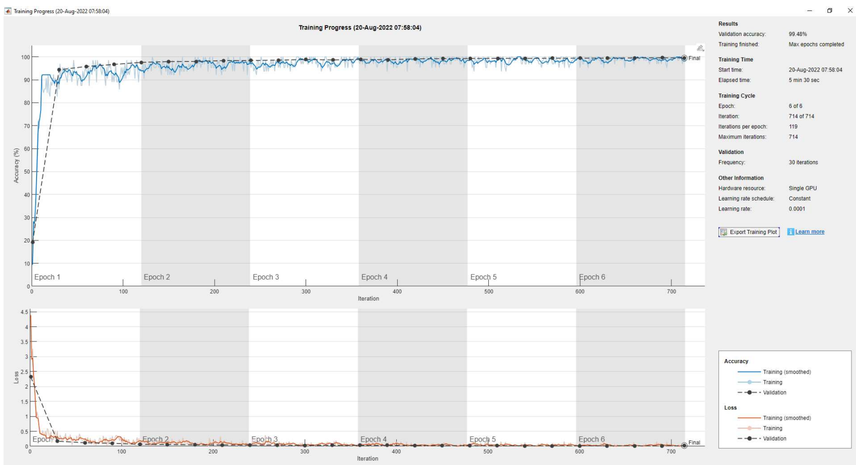

Figure 8. Training of the dataset with AlexNet

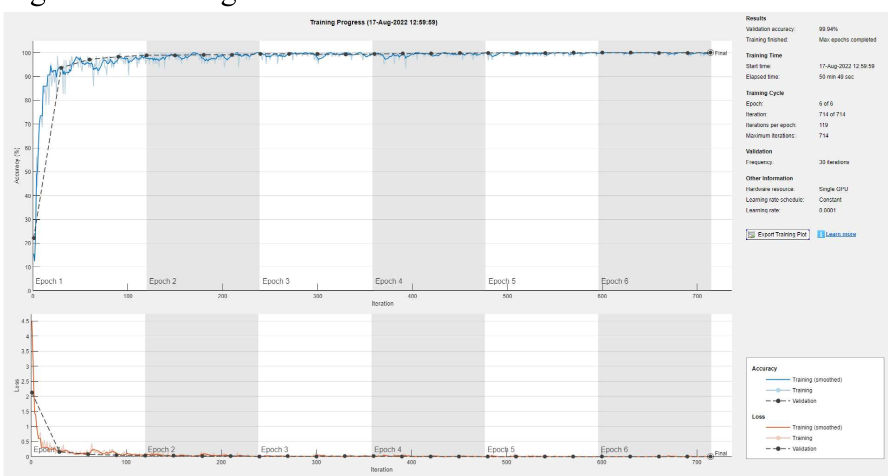

Figure 9. Training of the dataset with VGG19

Figure 10. Training of the dataset with GoogleNet

SEYBOLD Report
ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

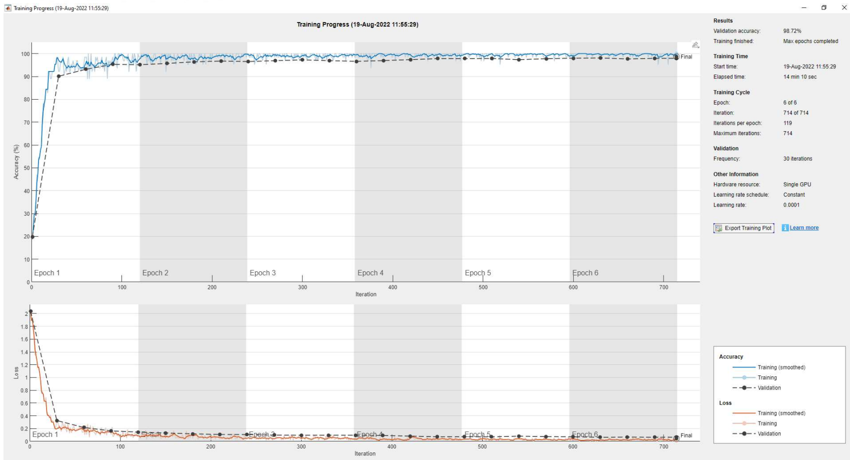

Figure 11. Training of the dataset with ResNet50

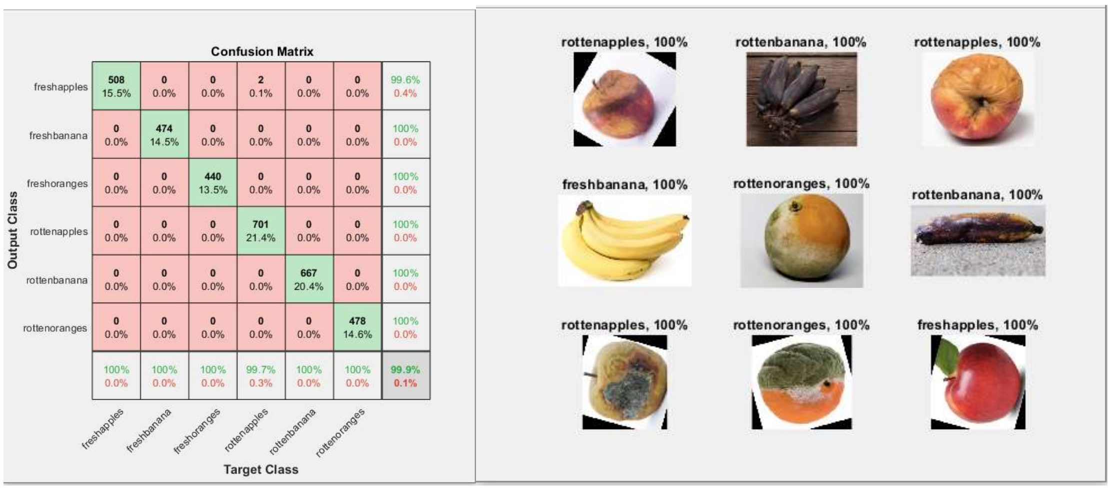

Figure 12 (a) Confusion Matrix of VGG19 (b) Sample Output Images

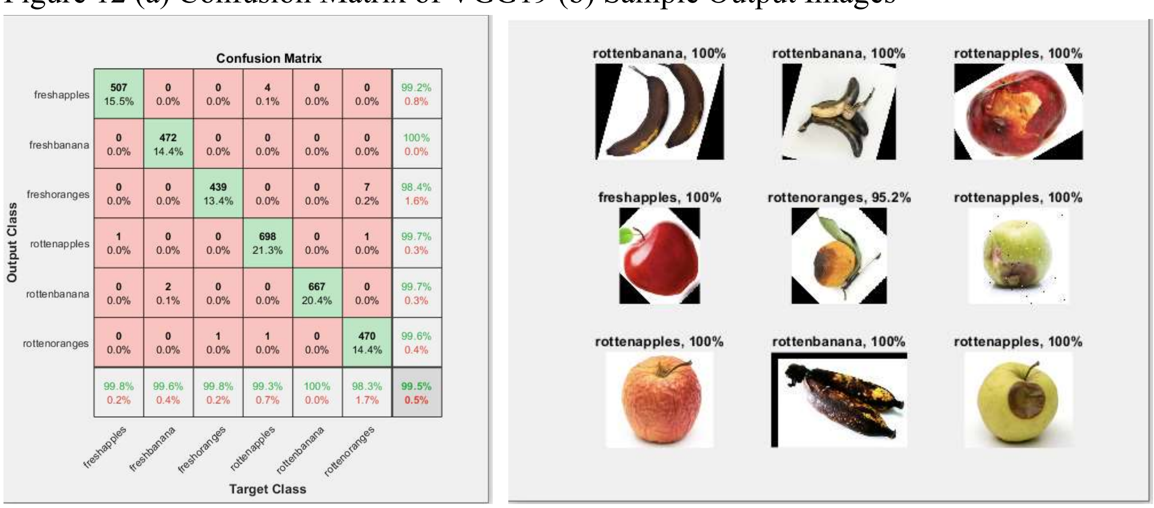

Figure 13 (a) Confusion Matrix of AlexNet (b) Sample Output Images

ISSN: 1533 - 9211
DOI: 10.5281/zenodo.7259996

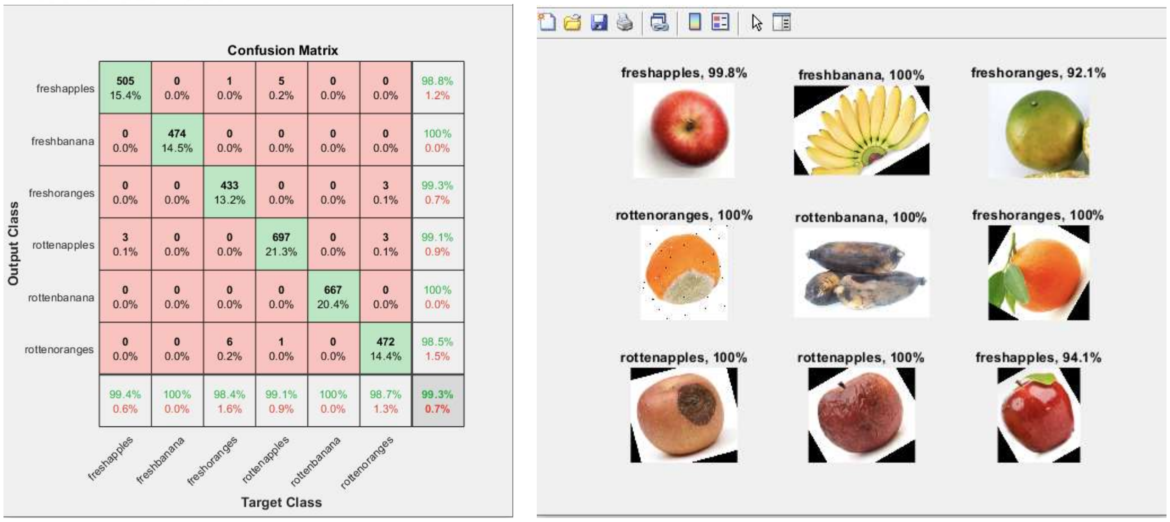

Figure 14 (a) Confusion Matrix of GoogleNet (b) Sample Output Images

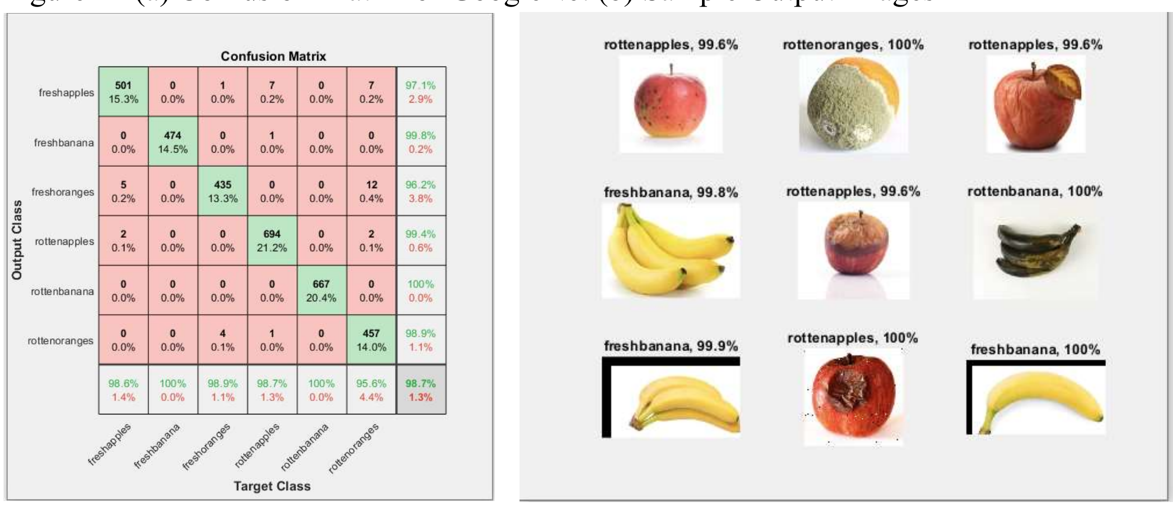

Figure 15 (a) Confusion Matrix of ResNet50 (b) Sample Output Images

Table 3. Results Obtained

S. No.	Model	Accuracy (%)	Time	Precision (%)	Recall (%)	F1 Score (%)
1	AlexNet	99.48	5 m 30 s	99.46	99.44	99.45
2	GoogleNet	99.33	8 m 39 s	99.28	99.30	99.29
3	VGG19	99.94	50 m 49 s	99.95	99.93	99.94
4	ResNet50	98.72	14 m 10 s	98.64	98.58	98.61

6 Discussion

We applied the transfer learning method to solve the issue of determining fruit quality. We have evaluated the different CNN models like AlexNet, VGG19, GoogleNet and ResNet50 with batchsize 64, learning rate 0.0001, 6 epochs and SGDM optimizer on the dataset to sort the fruits based on their quality. Fresh/rotten apples, fresh/rotten bananas, and fresh/rotten oranges were included in a dataset of six fruit groups to evaluate these pre-trained models. Each pre-trained model goes through the fine-tuning procedure to prepare it for training on the gathered

DOI: 10.5281/zenodo.7259996

dataset. The processing time and validation accuracy of these models are assessed. All the models have good validation accuracy. The Vgg19 model has the best validation accuracy at 99.94%, however, the processing time is excessively long at roughly 50 m 49 s. Both AlexNet and VGG19 work admirably, however, AlexNet outperforms VGG19 in terms of processing speed. As shown in Table 3 above, AlexNet has 99.48% validation accuracy and 5 m 30 s processing time. Again, the results above show us that, the best accuracy is obtained by VGG19 and the least is by ResNet50. According to the other metrics precision, recall and F1 score, VGG19 performs the best. We observed that AlexNet is the best architecture for detecting the quality of fruits.

7 Conclusion

We have evaluated the accuracy and processing time of four different CNN architectures, AlexNet, GoogleNet, VGG19, and ResNet50 and found the most effective method for distinguishing fresh and rotten fruits. The results indicate that the AlexNet model performs the best for sorting fruits and achieves state of art accuracy of 99.48% for a dataset with 64 batchsize, 0.0001 L.R. and 6 epochs with SGDM optimizer and the processing time was 5 m 30 s. Thus, AlexNet performs best in both time and accuracy.

Furthermore, we evaluated the models with other metrics like precision, recall, and F1 score. We concluded that the AlexNet architectures can be useful to the producers to improve their sorting process to detect fresh and rotten fruits.

As a future aspect, we can evaluate different CNN models to obtain the best accuracy and the least processing time.

References

[1] J. Ni, J. Gao, L. Deng, and Z. Han, "Monitoring the change process of banana freshness by googlenet," IEEE Access, vol. 8, pp. 228 369–228 376, 2020.

[2] M. Alencastre-Miranda, R. M. Johnson, and H. I. Krebs, "Convolutional neural networks and transfer learning for quality inspection of different sugarcane varieties," IEEE Transactions on Industrial Informatics, vol. 17, no. 2, pp. 787–794, 2020.

[3] O. A´ lvarez-Canchila, D. Arroyo-Pe´rez, A. Patinˇo-Saucedo, H. R. Gonz´alez, and A. Patino-Vanegas, "Colombian fruit and vegetables recognition using convolutional neural networks and transfer learning," in Journal of Physics: Conference Series, vol. 1547, no. 1. IOP Publishing, 2020, p. 012020.

[4] K. B. Mirra, R. Rajakumari, "Classification of Fruits using Deep Learning Algorithms", PREPRINT (Version1) Research Square, 01 April 2022, [https://doi.org/10.21203/rs.3.rs-1495878/v1]

[5] S. Behera, A. Rath and P. Sethy, "Maturity status classification of papaya fruits based on machine learning and transfer learning approach", Information Processing in Agriculture, vol. 8, no. 2, pp. 244-250, 2021.

[6] I. Hussain, S. Tan, W. Hussain, and A. Ali, "CNN transfer learning for automatic fruit recognition for future class of fruit," International Journal of Computer (IJC), vol. 39, no. 1, pp. 88–96, 2020.

[7] J. Pardede, B. Sitohang, S. Akbar, and M. L. Khodra, "Implementation of transfer learning

using vgg16 on fruit ripeness detection," Int. J. Intell. Syst. Appl, vol. 13, no. 2, pp. 52–61, 2021.
[8] S. Turaev, A. Abd Almisreb, and M. A. Saleh, "Application of transfer learning for fruits and vegetable quality assessment," in 2020 14th International Conference on Innovations in Information Technology (IIT). IEEE, 2020, pp. 7–12.
[9] S. A. P. N. Kavala and R. Pothuraju, "Detection of grape leaf disease using transfer learning methods: Vgg16 & vgg19," in 2022 6th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2022, pp. 1205–1208.
[10] R. Sustika, A. Subekti, H. F. Pardede, E. Suryawati, O. Mahendra, and S. Yuwana, "Evaluation of deep convolutional neural network architectures for strawberry quality inspection," Int. J. Eng. Technology, vol. 7, no. 4, pp. 75–80, 2018.
[11] https://www.kaggle.com/datasets/sriramr/fruits-fresh-and-rotten-for-classification
[12] X. Han, Y. Zhong, L. Cao, and L. Zhang, "Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification," Remote Sensing, vol. 9, no. 8, p. 848, 2017.
[13] J. Jaworek-Korjakowska, P. Kleczek, and M. Gorgon, "Melanoma thickness prediction based on convolutional neural network with vgg-19 model transfer learning," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019.
[14] B. Zhao, J. Feng, X. Wu, and S. Yan, "A survey on deep learning based fine-grained object classification and semantic segmentation", International Journal of Automation and Computing, vol. 14, no. 2, pp. 119–135, 2017.
[15] T. Ramu and A. Muthukumar, "A googlenet architecture based facial emotions recognition using EEG data for future applications," in 2022 International Conference on Computer Communication and Informatics (ICCCI), pp. 1–7, 2022.
[16] (He K., Zhang, Ren, & Sun, "Deep residual learning for image recognition", Proceedings of the IEEE conference on computer vision and pattern recognition, pp.770-778, 2016.
[17] https://jananisbabu.github.io/ResNet50_From_Scratch_Tensorflow/