E-fresh : Computer vision and IOT based system for food industry

Akshata Gawas
Dept. of Information Technology
A. P. Shah Institute of Technology
Thane, India
akshatagawas@apsit.edu.in

Krishita Tolia
Dept. of Information Technology
A. P. Shah Institute of Technology
Thane, India
krishitatolia@apsit.edu.in

Siddhesh Gaikwad
Dept. of Information Technology
A. P. Shah Institute of Technology
Thane, India
siddheshgaikwad@apsit.edu.in

Prof.Kiran Deshpande
Dept. of Information Technology
A. P. Shah Institute of Technology
Thane, India
Kbdeshpande@apsit.edu.in

Prof.Kaushiki Upadhyaya
Dept. of Information Technology
A. P. Shah Institute of Technology
Thane, India
Ksupadhyaya@apsit.edu.in

Prof.Sonal Jain
Dept. of Information Technology
A. P. Shah Institute of Technology
Thane, India
Sajain@apsit.edu.in

***Abstract*—The food industry is expanding every day and it is crucial to maintain the required standards which impact their market value. To maintain these standards, manpower is used which is inconsistent, expensive, and time-consuming. With the help of automation of classification, we can speed up this process with less expensive resources using Computer Vision along with the Internet of Things(IoT). The evolution of the Internet of Things(IoT) has played an important role in making the devices smarter and more connected. The large amount of data collected from these devices can be used for data analysis which helps the industries to plan their future decisions. This paper proposes the idea of implementing an infrastructure having a micro-controller that would accurately segregate three kinds of fruits into two categories i.e. Fresh and Rotten. The classification will be done with the help of the Deep Learning algorithm, Convolutional Neural Network(CNN) by using a dataset containing images of those three fruits and also considering the input from the sensors which include sensors such as alcohol sensor, methane sensor, etc. The infrastructure proposed in the paper considers the standards of Industry 4.0 which implies the real-world implementation of the infrastructure.**

***Index Terms*—Convolutional Neural Network(CNN), Internet of Things(IoT), Deep Learning(DL), Industry 4.0, Image Classification, Computer Vision**

I. Introduction

The fruit industry has become the third major after grain and vegetables. The rapid development of the fruit industry has brought us visible economic benefits, but also brought us a series of problems, fruit classification is one of them. Retail and supermarkets require manual labor to sort and classify fruits depending on their freshness. This includes not just labor costs, but also the time spent on these operations. Thus, to overcome this problem we need an automated system that can reduce the efforts of humans, reduce the cost of production and time of production. Many researches have been carried out which depicts the use of convolutional neural network(CNN) on various aspects [2] [3] [4] [14] [19] and many more. With the help of previous studies, we have proposed a model which will identify whether the fruit is fresh or rotten using a convolutional neural network(CNN). With the application of deep learning architecture in the field of image recognition, a convolutional neural network(CNN) known as one of the typical deep learning models, used for transfer learning has a good performance in image classification. CNN has various architectures which are developed with the evolving technologies, one of which is Inception-V3. It is a convolutional neural network architecture from the inception family that makes several improvements, considering accuracy as well as computational cost [9] [10] [11]. Thus, Inception-V3 is suitable for the fresh and rotten fruits dataset. Classification will take place based on color, shape, edge detection, ethanol emission, methane emission, and texture. The Fourth Industrial Revolution known as Industry 4.0 is the digital transformation of manufacturing and related industries and value creation processes. It involves automation of traditional manufacturing and industrial practices, using modern smart technology. Considering the standards of Industry 4.0 [5] [6] this project contains a conveyor belt along with the sensors such as methane, ethanol to detect spoilage of the fruit. Additionally, data analysis will provide information about the fruits which will in turn state their impact on the production.

II. Literature Review

Our literature review discusses the study of different classification algorithms used in past few years with their obtained accuracy. The among the recent ones, the study of Apple Fruit Detection and Maturity Status Classification [8] obtained

Year	Author	Problem Description	Dataset used	Algorithm	Learning Type	Accuracy	Limitations	Ref
2020	Sai Sudha Sonali Palakodati, Venkata RamiReddy Chirra	Fresh and Rotten Fruits Classification Using CNN and Transfer Learning	3 fruits: apples, bananas and oranges	CNN	Transfer Learning	97.82%	Small dataset with small number of convolution layers.	[14]
2015	Karen Simonyan, Andrew Zisserman	Very Deep Convolutional Networks for Large-Scale Image Recognition	ILSVRC-2012 dataset	CNN-VGG	Supervised Learning	top-5 test error: 6.8%	A fixed kernel size of 3x3 was used.	[15]
2020	Deepika Srinivasan, Mahmoud Yousef	Apple Fruit Detection and Maturity Status Classification	Kaggle's Fresh and Rotten fruits Image dataset	CNN-ResNet50	Supervised Learning	97.92%	Locally based system with small dataset.	[8]
2019	Mengying Shu	Deep learning for image classification on very small datasets using transfer learning	6000 images of dogs and cats.	CNN: VGGNet, GoogleNet, InceptionResNet	Transfer Learning	InceptionResNet: 96%, Inception V3: 95%	Problem of underfitting was observed.	[16]
2020	Yuhang Fu	Fruit Freshness Grading Using Deep Learning	6 fruits: apple, dragon fruit, kiwi, pear, banana, orange.	YOLO for detection and classification, CNN for freshness level regression	Supervised and Transfer Learning	91.49%	YOLOv3 localize objects through rectangular bounding boxes that results in background noises.	[17]
2019	Nguyen Truong Thinh, Nguyen Duc Thong	Mango Classification System Based on Machine Vision and Artificial Intelligence	Images of mangoes.	ANN	Unsupervised Learning	90%	ANN are more suitable for textual data rather images.	[1]
2013	Jagadeesh Devdas Pujari, Rajesh Yakkundimath	Grading and Classification of Anthracnose Fungal Disease of Fruits based on Statistical Texture Features	Images of fruits of type normal and affected by anthracnose	BPNN	Supervised Learning	Normal: 85.65%, Affected: 76.6%	Small dataset and BPNN is very sensitive to noisy data.	[18]
2018	Ling Zhu, Zhenbo Li, Chen Li, Jing Wu, Jun Yue	High performance vegetable classification from images based on AlexNet deep learning model	ImageNet	AlexNet	Transfer Learning	92.1%	Overfitting was observed as image dataset was small.	[19]
2019	Ranjit K N, Raghunandan K S, Naveen C, Chethan H K, Sunil C	Deep Features Based Approach for Fruit Disease Detection and Classification	Quadtree segmentation RGB images	CNN	Supervised Learning	93%	A fixed kernel size of 3x3 was used.	[20]

Literature Review

the highest accuracy of 97.92% by fine-tuning the ResNet50 architecture of the CNN algorithm. But it had a limitation as the system was locally based and the model was trained on a small dataset. In another paper on the topic of Fresh and Rotten Fruits Classification using CNN and transfer learning [1], various architectures of CNN were used to check the accuracy, and a simple CNN model was deployed having a small number of convolution layers, and the model was trained on a small dataset and an accuracy of 97.82% was obtained. A study on the Classification of images having dogs and cats [16] was reviewed, various CNN architectures like VGGNet, GoogLeNet., and InceptionResNet were used, and it was observed that InceptionResNet gave an accuracy of 96% while Inception V3 gave 95% accuracy. As the dataset was very small a problem of overfitting was observed in the study. A study on Mango Classification System based on Machine Vision and Artificial Intelligence [1] was studied where they used an ANN algorithm to sort mangoes and detect blemishes by considering features such as size, density, color, etc., and got an accuracy of 90%. A paper on the topic Deep Features Based Approach for Fruit Disease Detection and Classification [20] has described the use of the CNN algorithm on Quadtree segmented RGB images for detecting the fruit disease with an accuracy of 93%. However, a fixed- size kernel was used in the algorithm. A paper on the topic of High-performance vegetable classification from images based on the AlexNet deep learning model [19] was studied, the CNN algorithm AlexNet was used on the dataset ImageNet from which broccoli, pumpkin, cauliflower, mushrooms, and cucumber images were used and it gave an accuracy of 92.1% but as a small number of images were used, it may result into overfitting. So, we reviewed various kinds of algorithms and tried to overcome the limitations observed in the reviewed studies. Our dataset has an acceptable number of images, and we wish to expand it as we deploy the architecture. The algorithm selected is Inception V3 which is considered one of the efficient CNN algorithms as it has varied kernel size, batch normalization, and Auxiliary branches which will help to conquer the problem of vanishing gradient. The software part of our proposed system will be deployed on the cloud and the data obtained from the system will be used for data analysis.

III. Method

A. Architecture

The architecture consists of software as well as hardware. The spotlight of our project is to classify and segregate fruits according to their freshness. The paper proposes the implementation of Inception-V3 an advanced architecture of Convolutional Neural Network (CNN) using transfer learning [14] [16] [21] for image recognition and softmax function is used for image classification. For classification we need a supervised learning model which will brief us on whether the fruit is fresh or rotten so, Fig no. 2 shows the whole process to be accomplished. Firstly a predefined dataset stored in the database will be passed for preprocessing that mainly includes resizing, noise reduction, etc. and it is an important step to be performed before passing it for training. After that, the images will be passed to the model for training, and after accomplishing it, the trained model will be saved in the database having the desired accuracy. For calculating accuracy, we pass the test set for prediction and check whether the predicted values match the original value.

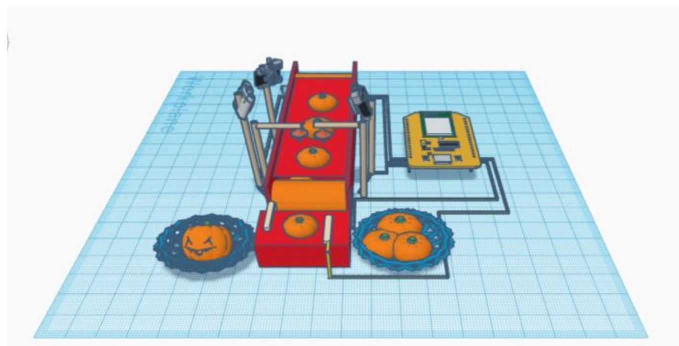

Fig. 1. Demonstration Model

Now we shall have a look over the hardware section (Segregation section) i.e. Fig no.1. The hardware comprises of one conveyor belt. The conveyor belt is built considering the Industry 4.0 standards. Sensors, nodeMCU8266, camera modules, servo motors are incorporated in the conveyor belt. MQ3 and MQ4 sensors are used to detect the concentration of ethanol [23] and methane gas emitted by the fruits, three camera modules are placed at suitable positions to capture images of the fruits that are kept on the conveyor belt to classify and segregate according to their freshness, nodeMCU8266 is used to pass images to the trained model which is stored on the cloud, the servo motor is used to push the rotten fruits into a separate basket. In nutshell, the fruits will be carried forward through the conveyor belt, recognition and classification will be done using the trained model, Google Colab, Keras, TensorFlow, NumPy, etc then the fresh fruit will fall in the fresh fruit basket and rotten fruits will be pushed into a separate basket. After the whole process analytical graphs will be presented for future use on the web portal. Colour, size, ethanol level, methane level, texture, and edge detection are the parameters which will be used for data analysis. All the acquired data from various sensors and the prediction from classifying algorithm will help to segregate the fruits on the basis of their freshness.Various papers depicts the use of cloud computing which has resulted into increasing efficiency of the model [22]. So, our proposed framework also intends to store all the collected data on a cloud which will be used in data analysis which will help in taking better decisions considering the business point of view.

B. Dataset and Data Augmentation

The dataset used for training the model is downloaded from Kaggle and it consists of images of fresh apples, bananas, and oranges, and rotten apples, bananas, and oranges i.e. it consists of 6 classes. The total number of images in the dataset are

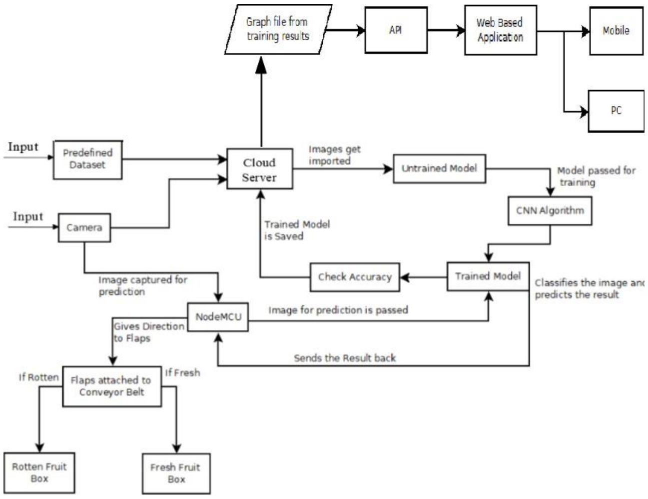

Fig. 2. Demonstration of proposed architecture

13,599 where 10901 images are used for training and 2698 for testing purpose. Some of them are displayed in Fig. 3. Data augmentation was performed by rescaling the image, applying shear, applying horizontal flip, and zooming into the image. Then the entire dataset is shaped to 224x224x3 and into a NumPy array for an easier convolution process.

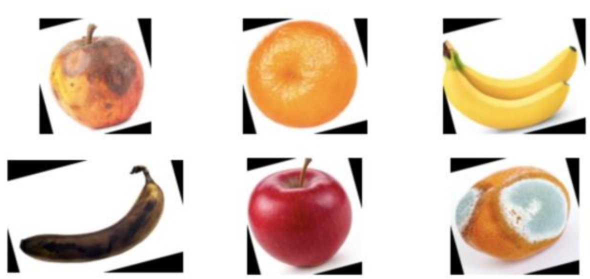

Fig. 3. Some images from dataset

C. Convolutional Neural Network (CNN)

A convolutional neural network [2] is a category of deep neural networks, widely used to analyze visual images. CNN pulls out an image feature and converts it to a lower level in terms of dimension. CNN follows a sequence model that works to build a network and ultimately provides a fully connected layer where all the sensors are connected and the result is processed. CNN is used for image classification and recognition due to its high accuracy. Unlike other algorithms with pixel vector CNN prevents a lot of spatial interaction between pixels, it uses adjacent pixel information to effectively downsample the image first by convolution and then uses a prediction layer at the end. In fig no. 4, the initial size of the image is 224 x 224 x 3. Proceeding without convolution will generate 1,00,352 neurons, but after applying convolution the input tensor dimension will reduce to 1 x 1 x 1000 generating 1000 neurons in the first layer. The convolution layer is the first layer, it is used to extract features from the input image. Kernel/Filter is the element responsible for carrying out the convolution operation. The mathematical function of convolution is performed between the input image and the filter. By moving the filter above the input image, the dot product is taken between the filter and parts of the input image concerning the size of the filter. The output is called the Feature Map which gives us image information such as corners and edges. [3] proposes a framework of six layer convolution layers and 5 pooling layers. Pooling layer: The main purpose of this layer is to reduce the spatial volume of the integrated feature map. This is to reduce the calculation costs required to process data by dimensionality reduction. There are two types of Pooling Max Pooling and Average Pooling. Max Pooling returns the highest value on the part of the image covered by Kernel. On the other hand, Average Pooling returns the average of all values from the part of the image covered by Kernel.

Fully Connected Layer: A fully Connected Layer consists of weights, biases, and neurons and is used to connect neurons between two different layers. These layers are usually placed before the output layer. It is used to classify images between different categories by training. Activation Layer : It is one of the most important parameters of the CNN model activation function. This layer is the last layer of CNN. It sits at the end of the FC layer. It decides which model information should shoot forward and which should not end up in the network. Generally Logistic is used for binary classification whereas sigmoid and softmax are used for multi-classification.

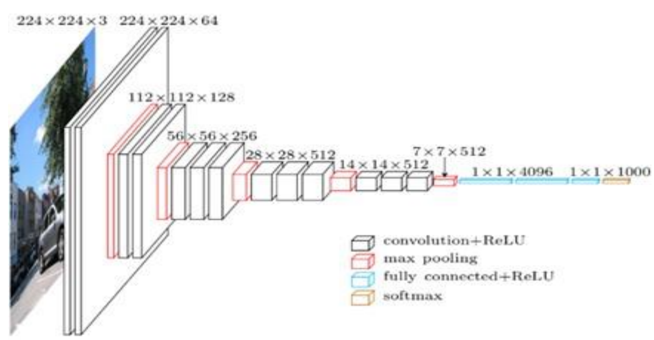

Fig. 4. Overview of Convolutional Neural Network

D. Network Architecture

As Jeff Tang suggests in his book Intelligent Mobile Projects with TensorFlow, Inception V3 is a convolutional neural network for assisting in image analysis and object detection and got its start as a module for GoogLeNet. It is a 48-layer deep neural network with 24M parameters and has greater than 78.1% accuracy on the ImageNet dataset. Inception v3 optimizes the process by including techniques like factorized convolutions, dimension reduction, regularization, label smoothing, and batch normalization. Neural networks like VGGNet use a fixed size of the kernel while Inception V3 uses different sizes of kernels to extract different sizes of features optimally. The study has suggested that very deep neural networks are prone to overfitting and even face the problem of vanishing gradient. So, Inception V3 overcomes these problems by introducing a bit wider network rather than deeper. [11] suggests that with large dataset, the accuracy of CNN increases which in turn helps us in better classification. Fig. no.5 shows the naive block of Inception which with evolving technology has become efficient to form Inception V3 with 3 blocks A,B and C.

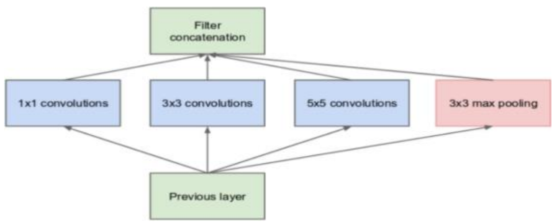

Fig. 5. Naive block of Inception[10]

Inception V3 introduces 3 different blocks which when applied between the convolution layers gives a computational as well as memory optimum solution. [10] suggests that any reduction in computational cost results in a reduced number of parameters. This means that with suitable factorization, we can end up with more disentangled parameters and therefore with faster training. Convolutions like 5x5 or 7x7 tend to be expensive than 3x3 or 1x1 convolutions, so the blocks A, B, and C introduced are shown in figure no.6.

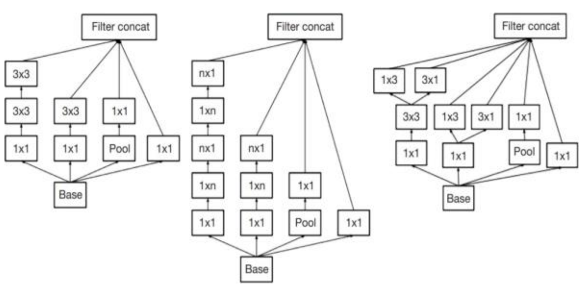

Fig. 6. Block A, block B, block C of Inception-V3 respectively [11]

In block A, factorization is introduced. Consider 5x5 and 3x3 convolution, if we evaluate 5x5 we get 25 and 3x3 two times we get 18, so in block A, 2 3x3 convolutions were introduced to reduce the dimensionality of the input channels. In block B, they further used the factorization as, if we consider nxn convolution it was factorized as 1xn and nx1, this approach was found out to be 33% cheaper than the single nxn. In block C, the architecture was made wider instead of deeper. If it was made deeper, there were chances of image distortion and loss of information. Batch Normalization as proposed in paper [12], is a method used to make artificial neural networks faster and more stable through normalization of the layers' inputs by re-centering and re-scaling. Batch Normalization was introduced in the auxiliary classifiers which helps in the process of regularization. Label smoothing was also introduced in inception V3 to prevent it from predicting the labels too confidently that is to overcome the problem of overfitting.

IV. Evaluation

A. Study of Sensors

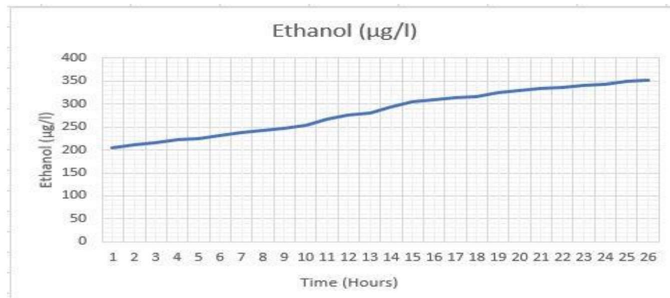

Fig. 7. Ethanol levels

Ethanol is a naturally occurring compound caused by the fermentation of fruit sugar [23]. It is a simple alcohol with the chemical formula C2H6O. Fig.no. 7 shows ethanol production in banana with the passage of time . Ethanol production begins as bananas begin to ripen with the appearance of black spots on the banana, which causes the fruit to soften. Over time, the black spots increase, followed by a period of ripening when an additional form of fungal growth appears on the fruit's surface.

B. Experiments and results

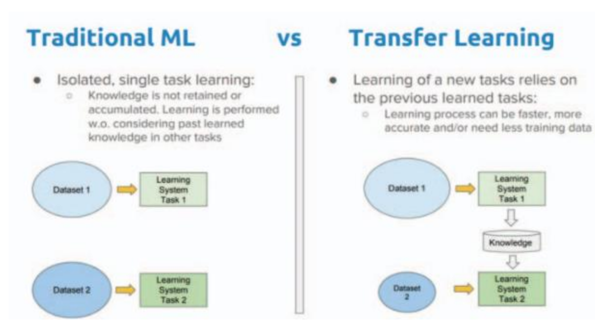

Fig. 8. Difference between tradition CNN model and Transfer Learning model[21]

Transfer Learning [21] is a method of taking a pre-trained model on some dataset and applying it to another dataset with some necessary changes. It is like learning techniques of understanding a concept from certain information and applying that learning to other information. The major difference between traditional CNN and transfer learning model is depicted in Fig no.8. There are various types of CNN architecture which include VGGnet, ResNet, etc. but our system uses the CNN's Inception V3 model trained on the dataset ImageNet as the Inception model not just goes deeper but also widens the neural network architecture. As we used transfer learning, the number of trainable parameters reduced from 21,853,985 to 51,201. The proposed model, with the help of fine-tuning provided a training accuracy of 98.33 and validation accuracy of 99.42%. The Fig. no.9 show the graph of accuracy of training and validation set throughout the training of model, while Fig no.10 shows the loss incurred to training and validation set during the training of model

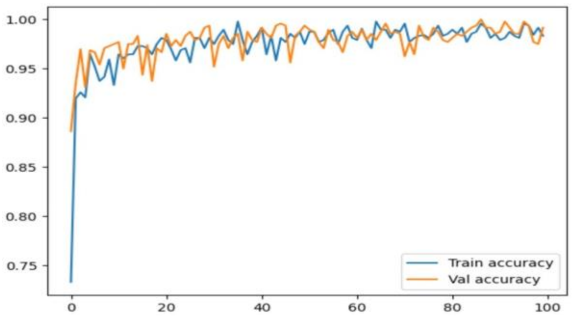

Fig. 9. Accuracy of training and validation set

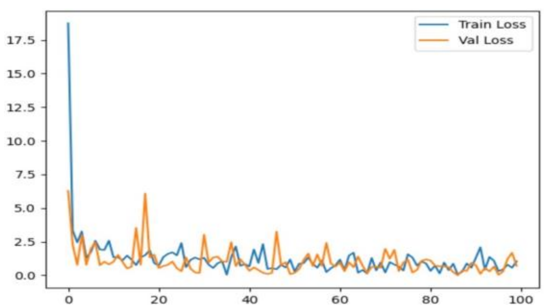

Fig. 10. Loss of training and validation set

Batch-size: Batch size is the number of input images provided to the neural network. It plays a crucial role in determining accuracy. It is considered as larger the batch size, more time is taken for training the model, resulting in a decrease in accuracy. We trained our model with batch sizes of 8,16,32 and 64. As we can see in the graph fig no.11, the accuracy increased from batch-size 8 to 16, while there was a gradual decrease in the accuracy from 16-64, here the optimum batch size is 16.

Epochs Now, let's see the effect of the number of epochs on the accuracy of the model. Epochs are nothing but the number of iterations. Now by keeping the batch size constant at 16, using the RMSprop optimizer, the model was trained as 5,10,20,30,40,50,60,70,80,90,100 epochs. The graph fig no.12 shows that at the 100th epoch the accuracy noted was 99.17%. We can see that as we increase the number of the epoch, the accuracy increases reaching its optimum value.

Confusion Matrix summarizes the performance of the classification algorithm. It gives a better idea of how the algorithm is performing in making predictions and various types of errors can be observed. Fig no.13 shows the Confusion Matrix of our trained model.

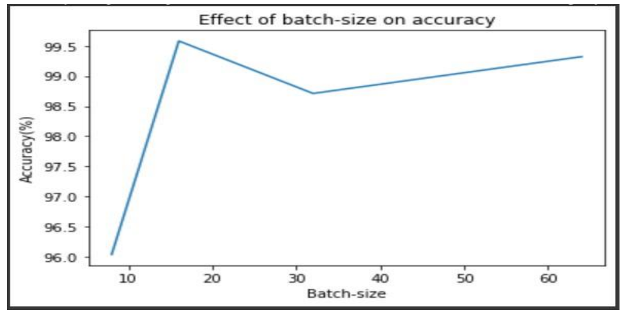

Fig. 11. Effect of batch-size on accuracy

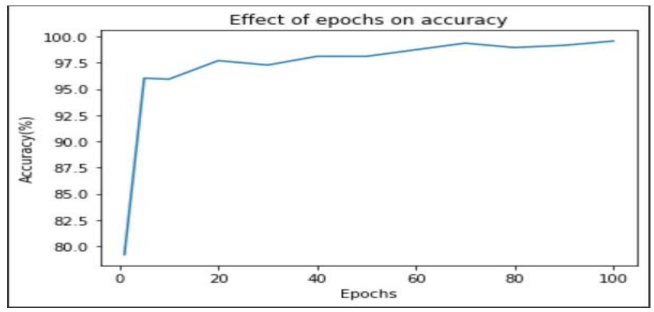

Fig. 12. Effect of epochs on accuracy

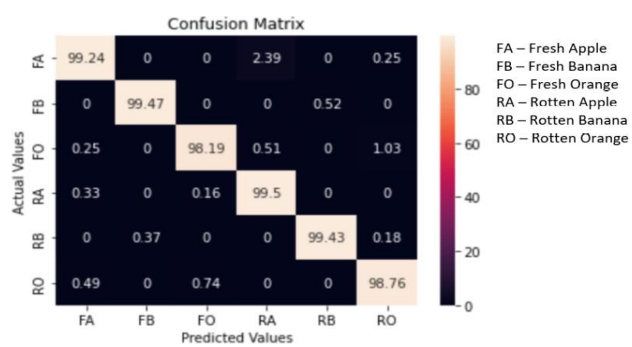

Fig. 13. Confusion Matrix

V. Conclusion

The process of recognizing, classifying, and segregating fruits according to their freshness is proposed in this paper. CNN architecture i.e., Inception-V3 using transfer learning is implemented in this proposed framework. Datasets with different fruits such as apple, banana, orange are used to train and test the model. The accuracy of the trained model with epochs 100 and batch-size 16 was recorded as 99.17. A touch of the Internet of things (IoT) is given to the architecture by using sensors, camera modules and sending data over the internet to the trained model stored on a cloud. Sensors give an additional advantage along with image classification to filter fresh and rotten fruits. While designing the hardware model the standards of Industry 4.0 were kept in consideration. This proposed framework would be an aid for the food industries as a huge amount of capital and time is spent on labor-intensive and repetitive tasks. So to conquer the drawbacks, this proposed framework can be implemented. Data analysis will provide information about the fruits which will, in turn, state their impact on the production. Further, the analytical report will be prepared on the web portal.

References

[1] Nguyen Truong Thinh, Nguyen Duc Thong, and Huynh Thanh Cong, "Mango Classification System Based on Machine Vision and Artificial Intelligence," Ho Chi Minh City University of Technology, VNU HCMC, 2019 IEEE 7th International Conference on Control.

[2] Rajesh Yamparala, Ramaiah Challa, V.Kantharao, P.Seetha Rama Krishna, "Computerized Classification of Fruits using Convolution Neural Network", IEEE 7th International Conference on Smart Structures and Systems ICSSS 2020.

[3] M. Shamim Hossain , Senior Member, IEEE, Muneer Al-Hammadi and Ghulam Muhammad , Member, IEEE, "Automatic Fruit Classification Using Deep Learning for Industrial Applications," IEEE transactions on industrial informatics, vol. 15, no. 2, February 2019.

[4] Liuchen Wu, Hui Zhang, Ruibo Chen, Junfei Yi, "Fruit Classification using Convolutional Neural Network via Adjust Parameter and Data Enhancement", 12th International Conference on Advanced Computational Intelligence (ICACI).

[5] Ms. Jayshree C. Phad, Ms.Kajal A. Nagare, Ms. Shubhangi R. Kalokhe, Prof. A. R. Khandare, Prof. D. Y. Thorat5, "Conveyor Belt Fault Detection Using IOT", 2019 IJRTI — Volume 4, Issue 3 — ISSN: 2456-3315.

[6] G. Lodewijks, W. Li, Y. Pang, X. Jiang, "An Application of the IoT in Belt Conveyor Systems".

[7] Anuja Bhargava, Atul Bansal "Fruits and vegetables quality evaluation using computer vision: A review", Journal of King Saud University – Computer and Information Sciences(2018).

[8] Deepika Srinivasan, Mahmoud Yousef, "Apple Fruit Detection and Maturity Status Classification", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-9 Issue-2, July 2020.

[9] CHENG WANG, DELEI CHEN, LIN HAO, XUEBO LIU, YU ZENG, JIANWEI CHEN, AND GUOKAI ZHANG, "Pulmonary Image Classification Based on Inception-v3 Transfer Learning Model", Digital Object Identifier 10.1109/ACCESS.2019.2946000.

[10] Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, "Rethinking the Inception Architecture for Computer Vision", arXiv:1512.00567v3 [cs.CV] 11 Dec 2015.

[11] Qing Guan, Xiaochun Wan, Hongtao Lu, Bo Ping, Duanshu Li1, Li Wang, Yongxue Zhu, Yunjun Wang1, Jun Xia,"Deep convolutional neural network Inception-v3 model for differential diagnosing of lymph node in cytological images: a pilot study", Ann Transl Med 2019;7(14):307 — http://dx.doi.org/10.21037/atm.2019.06.29

[12] Sergey Ioffe, Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", arXiv:1502.03167 11 Feb 2015.

[13] Christian Szegedy, Wei Liu, "Going deeper with convolutions", arXiv:1409.4842v1 [cs.CV] 17 Sep 2014

[14] Sai Sudha Sonali Palakodati, Venkata RamiReddy Chirra,"Fresh and Rotten Fruits ClassificationUsing CNN and Transfer Learning"

[15] Karen Simonyam, Andrew Zisserman,"Very Deep Convolutional Networks for Large-Scale Image Recognition"

[16] Mengying Shu,"Deep learning for image classificationon very small datasets using transfer learning"

[17] Yuhang Fu,"Fruit Freshness Grading Using Deep Learning"

[18] Jagadeesh Devdas Pujari, Rajesh Yakkundimath, "Grading and Classification of Anthracnose Fungal Disease of Fruits based on Statistical Texture Features"

[19] Ling ZhuZhenbo Li, Chen Li, Jing Wu, Jun Yue, "High performance vegetable classification from images based on AlexNet deep learning model"

[20] Ranjit K N, Raghunandan K S, Naveen C, Chethan H K, Sunil C, "Deep Features Based Approach for Fruit Disease Detection and Classification"

[21] Ribani, Ricardo and Marengoni, Mauricio, "A survey of transfer learning for convolutional neural networks",2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T)

[22] Channe, Hemlata and Kothari, Sukhesh and Kadam, Dipali, "Multidisciplinary model for smart agriculture using internet-of-things (IoT), sensors, cloud-computing, mobile-computing & big-data analysis", Int. J. Computer Technology & Applications

[23] Naveed Shahzad, Usman Khalid, Atif Iqbal, Meezan-Ur-Rahman, "eFresh – a Device to Detect Food Freshness", International Journal of Soft Computing and Engineering (IJSCE)