Identification of Fruits Using Deep Learning Approach

To cite this article: Deepali M Bongulwar 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1049** 012004

IOP Conf. Series: Materials Science and Engineering 1049 (2021) 012004 doi:10.1088/1757-899X/1049/1/012004

Identification of Fruits Using Deep Learning Approach

Deepali M Bongulwar
Department of Electronics and Communication, Sri Satya Sai University of Technology and Medical Science, Sehore (MP), India,Email:ddeepali2006@yahoo.com

Abstract: The paper intends to build a Model for the identification and classification of fruits using the concept of deep learning. The objective is to build an automatic system for feature extraction using convolutional neural networks. The system can sort the fruits. It can be put-to-use in checking the condition of fruits, whether they are fresh or not. The self-service system in the retail market can use it to identify fruits. The proposed system uses high quality 'ImageNet' dataset. The dataset consists of five different categories of fruit images. Dataset is very challenging. The model uses Convolutional Neural Networks to identify fruits from images. The accuracy obtained is 92.23%. Deep learning outperforms machine learning algorithms.

1. Introduction

There are various applications of identification and classification of fruits. Mostly, it is used in the identification of fruits in supermarkets and the robotic harvest in the agricultural field. It can be adopted in the identification of plant disease and species in the agricultural domain. In the supermarket, the items are identified physically by the sales associate, or through the self-service systems. The online retailer, Amazon built up a chain of grocery 'Amazon Go Grocery', which empowered the customer a just walk out shopping. The customers' can enter the store with the free app 'Amazon Pay' on their smartphone and shop like other supermarkets. The customer should have an Amazon account. The supermarket has an enormous number of digital cameras and sensors. Utilizing the concept of deep learning and computer vision algorithms, the technology recognizes the items the customers' pick. You need not stand in a line or checkout. The charges of picked items are deducted from the account of the customer [1].This paper deals with the development of the identification system. An image classifier is trained to identify different images of fruits. The motivation behind the use of this technology is to enhance the identification process done by the self-service systems in the supermarket. Here, explicitly the system depends on the convolutional neural networks (CNNs).CNNs have a large number of neural network layers that can progressively learn different features [2].CNNs are accustomed either to collect or count the fruits. The models using deep learning have mostly employed a previous couple of years for autonomous feature extraction. These models are most powerful in feature extraction. [3].They have obtained exceptional accuracy in several tasks like the classification of images, recognition of objects, and speech recognition [4] [5].A dataset of five categories, constituting 4,760 images is employed to check the functioning of the model. The greatest accuracy of classification achieved on the dataset is 92.23%.

The first part of the paper explains the objective of the paper, its need and the related work. Secondly, it describes the achievements in deep learning followed by the methodology. It describes the structure of CNN Model used and the training dataset as well as the obtained performance. Finally, it ends with the conclusion and future scope.

2. Deep Learning

Deep learning specifies that a computer figures out how to perform classification from images without manually extracting the features [2]. The deep learning models have accomplished the state-of-art accuracy beyond human-level capabilities. The training of deep learning models is done by utilizing a

huge amount of labeled data. Convolutional neural network architecture constitute numerous layers. The word 'deep' is used to refer more number of hidden layers. The deep learning is similar to machine learning that utilizes numerous layers with nonlinear processing units [6] [7].
The features extracted from every layer are applied as input to the succeeding layer. CNNs are categorized as the deep learning algorithms. These networks contain many convolutional layers and some fully connected layers. They additionally utilize pooling layers. The different deep learning architectures (DLA) like Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN) and Deep Belief networks (DBF) show state-of-the-art results in numerous areas. Autoencoders can be used for generating handwriting/speech/ images. Deep Belief Networks use Restricted Boltzmann Machine (RMB). (RMBs)/autoencoders can be used to train more deep networks. CNNs gives exclusively best accomplishment in object recognition. This made us choose deep learning to identify fruits from images. Deep neural networks perform much better than any machine learning algorithms. Furthermore, deep neural networks, particularly convolutional neural networks have demonstrated to get extraordinary outcomes in the recognition of images [8].

3. Methodology

This section outlines the methodology and dataset used to develop the Fruit Identification System and presents the system that has been used.

3.1. Convolutional Neural Network

Here, a convolutional neural network has been used. It consists of convolutional layers, batch normalization layers, pooling layers, ReLU layers and fully connected layers. Convolutional layers comprise kernels of different size. The depth of the kernels is similar to the given image. The neurons of the kernel are connected only to the local regions of the input image called the receptive field. Every neuron has its own weights. The kernels slide horizontally and vertically on the input image and result in a two-dimensional feature map. The parameter, stride decides how a kernel can slide on the input images. Pooling layers turn down the spatial dimensions of the input images and lessen the number of calculations. Thus, pooling layers control the overfitting of the Model. Here, the pooling layers have filters of size 2 x 2 and a stride of 2.The size of the image is changed to one-fourth of the input. ReLU (Rectified Linear Unit) layer is the activation function, max (0, x).These layers make no changes in the size of the network, but increase its nonlinearity. Finally, neural networks have fully connected layers. Every neuron of a fully connected layer is connected to the resulting output of the past layer [8].
This is generally the outermost layer of the network. Softmax layer is utilized for anticipating a category from numerous categories. The input comprises images of size 224x224x3 pixels. The MATLAB software with Neural Network Toolbox is utilized to create this convolutional neural network.

3.2. The Architecture of the proposed CNN Model

The proposed framework uses a deep learning architecture. The model has 41 CNN layers. The figure 1 below shows the architecture of the CNN Model. It comprises feature extraction and classification. The input images are cropped to remove any unwanted information. All the images are resized to 224X224X3 (3 corresponds to three color channels Red, Green, and Blue).The feature extraction is done by applying convolution and pooling layers number of times. The model shows six consecutive blocks and fully connected layers. First two blocks contain two consecutive convolution layers and max pooling layers and later on four consecutive layers of convolution. Each block has rectified linear unit (ReLU).This is non-linear activation function [9]. Batch normalization layers are applied after convolutional layers and before ReLU layers. This decreases the training time and reduces the sensitivity to network initialization. Here, the size of filters is fixed i.e. 3×3 for convolution and 2×2 for max-pooling are utilized. The CNN layers shows the feature maps of the images at different levels. The last layer is the classification layer. Here, the two-dimensional feature maps are flattened into one-dimensional feature vector. This resulting feature vector is applied as input to the outermost layer. The outermost layer is a fully connected layer of neurons with the number of categories equal to the number of input categories. The output layer uses softmax function. It finds the probability for each category and outputs the class with maximum probability. The dropout layers are added before the fully connected layer. There are four dropout layers. The dropout layers are used for regularization of the network and thus avoids overfitting.

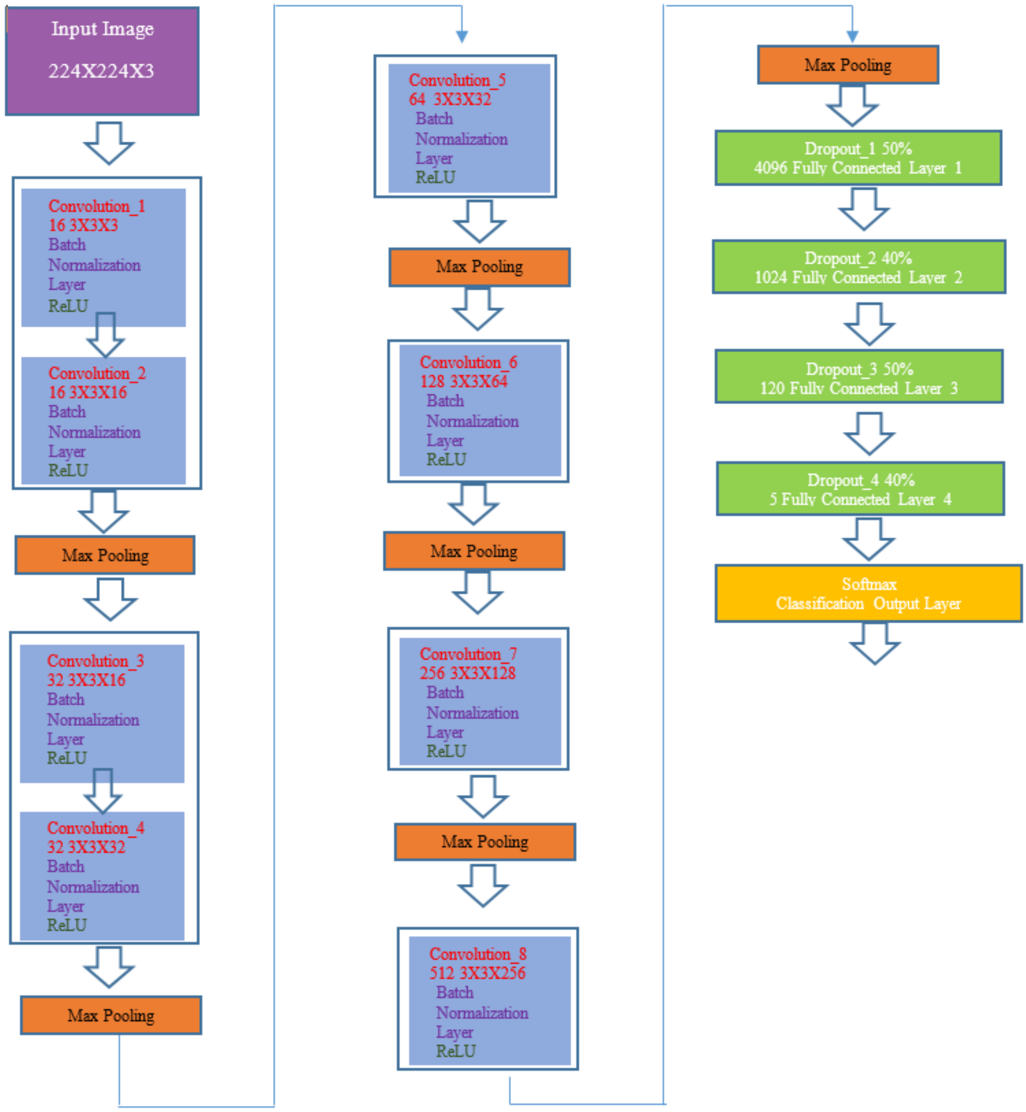

Figure 1. The Architecture of the proposed CNN Model

3.3. Dataset

All the images are downloaded from the 'ImageNet', free open dataset. Dataset consists of images of five different categories as shown below in table 1. The five categories are Apple, Banana, Grapes, Litchi and Mango. The RGB images, with three color channels R, G, and B are utilized in the dataset. The randomly selected sixteen images of the used dataset are as shown in figure 2.The total number of images including all the five classes in the dataset are 4,760 as shown in table 1.

The images in the dataset contain fruits of the same class with a different size. The background of the images is not homogeneous. The dataset has different poses of the same types of fruits. Fruits are with different positions and views like top view, side view, different background, half cut, sliced on the dish, cut in pieces, half-eaten, showing the seed, partially occluded. Sometimes fruits are fresh, sometimes rotten, or in bundles. Some images are with poor illumination, with different light effects, covered with snow or net, decorated, painted, along with leaf, on trees, inside a box or a plastic bag etc.

Black/white images are also considered. The dataset is very challenging. This is a very important factor in the proposed system.

Table 1. Input dataset

Label	Count
Apple	977
Banana	918
Grape	995
Litchi	964
Mango	906

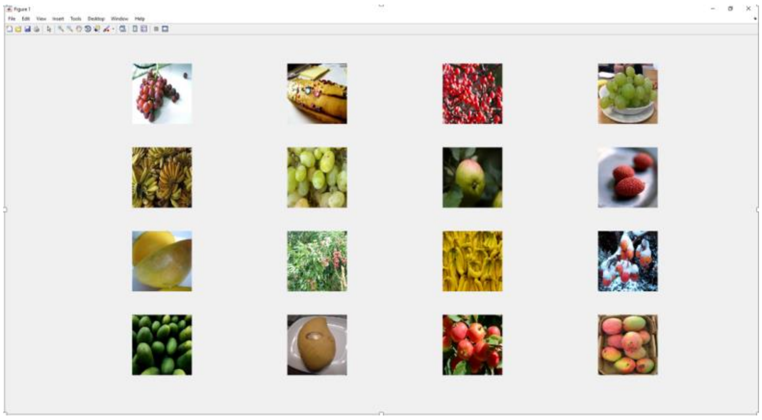

Figure 2. Sample Input Images

4. Experimental Results and Discussion

The images in the dataset are preprocessed and resized to 224x224x3 and cropped to remove any additional information. The dataset is divided into training and validation datasets in which 90% of the images are trained and 10% are validated. The optimizer, stochastic gradient descent (SGD) with piecewise learning is used. The minibatch size is 128. The initial learning rate is 0.01. The learning rate drop factor is 0.02. The learning rate drop period is 40 epochs. The images in the training dataset are augmented. The 128 samples in every minibatch are augmented as below:

-Randomly rotated between 0^0 and 90^0
-Horizontal reflection
-Horizontal and vertical translation

The proposed CNN model has achieved classification accuracy of 92.23%. The training and validation progress for each iteration of the epoch is shown graphically in figure 3. The labels on the output images show that they are accurately identified. The output images in figure 4 are pieces of fruits, still they are perfectly identified. The model is robust and gives very good accuracy.

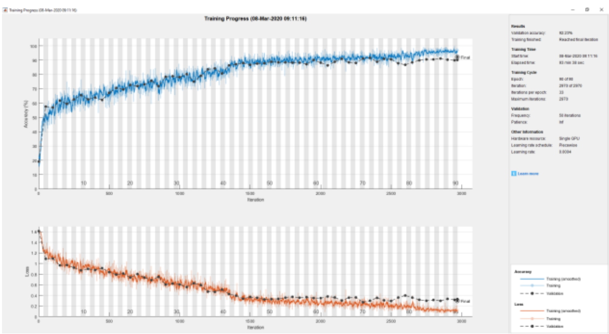

Figure 3. Training Progress versus Epoch Number

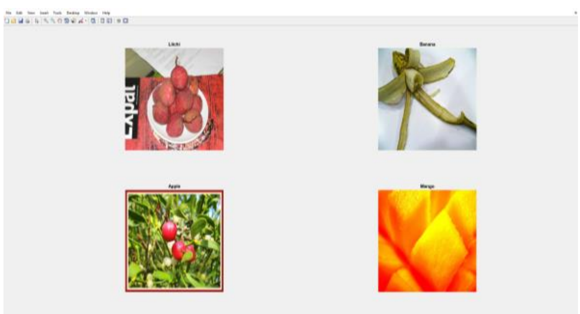

Figure 4. Resulting Output Labeled Images

5. Conclusion

The deep learning model for fruit identification and classification is developed in the paper. The paper has presented a system that develops an autonomous identification of fruits by the self-service system in the supermarket. The CNN Model has accomplished outstanding accuracy on the dataset.

As a future scope, the model can be used to train more variety of fruits. It can also examine the impact of various parameters like activation function, pooling function and optimizers.

References

[1] Frida Femling, Adam Olsson and Fernando Alonso-Fernandez 2018 Fruit and Vegetable Identification Using Machine Learning for Retail Applications *14th Int. Conf. on Signal-Image Technology & Internet-Based Systems (SITIS)* pp 9-15
[2] Y LeCun, L Bottou, Y Bengio and P Haffner 1998 Gradient-based learning applied to document recognition *Proc. of the IEEE* vol 86 pp 2278-2324
[3] X Yuan et al 2018 Deep Learning-Based Feature Representation and Its Application for Soft Sensor Modeling with Variable-Wise Weighted SAE *IEEE Transactions on Industrial Informatics* **14** pp 3235-43
[4] Z Gao, L Wang, L Zhou and J Zhang 2017 HEp-2 Cell Image Classification with Deep Convolutional Neural Networks *IEEE J. Biomed. Heal. Informatics* vol 21 pp 416-28
[5] T Kooi et al 2017 Large scale deep learning for computer-aided detection of mammographic lesions *Med. Image Anal* **35** pp 303–12
[6] Stuart J Russell, Peter Norvig 1995 *Artificial Intelligence-A Modern Approach* ed Egnlewood Cliffs (New Jersey: Simon and Schuster Company and Prentice-Hall) pp 79-80.
[7] Jurgen Schmidhuber 2015 Deep learning in neural networks: An overview of Neural Networks *National Centre for Biotechnology Information* **61** pp 85-117
[8] Horea Mureşan, Mihai Oltean 2018 Fruit recognition from images using deep learning *Acta Universitatis Sapientiae Informatica* **10** pp 26-42
[9] M Shamim Hossain, Muneer Al-Hammadi and Ghulam Muhammad 2018 Automatic Fruit Classification Using Deep Learning for Industrial Applications *IEEE Transactions on Industrial Informatics* **15** pp 1027-34.
[10] G E Dahl, T N Sainath and G E Hinton 2013 Improving deep neural networks for LVCSR using rectified linear units and dropout *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* (Vancouver BC Canada) pp 26-31
[11] Luis Santos Filipe N Santos, Paulo Moura Oliveira and Pranjali Shinde 2019 Deep Learning applications in agriculture: a short review *Robotics conf.* (Springer) pp 139-51
[12] Inkyu Sa, Zongyuan Ge, Feras Dayoub, Ben Upcroft, Tristan Perez and Chris McCool *DeepFruits:* A Fruit Detection System Using Deep Neural Networks
[13] Dan Claudiu Cire¸ Ueli Meier, Jonathan Masci, Luca Maria Gambardella and Jurgen Schmidhuber 2011 Flexible, high performance convolutional neural networks for image classification *Twenty-Second Int. Joint Conf. on Artificial Intelligence* pp 1237-42
[14] Rupesh K Srivastava, Klaus Greff and Jurgen Schmidhuber 2015 Training very deep networks, Advances in neural information processing systems *Twenty-Eight Int. Conf. on Neural Inf. Processing Systems* (Montreal Canada) pp 2377-85
[15] Woo Chaw Seng and Seyed Hadi Mirisaee 2009 A New Method for Fruits Recognition System *IEEE Int. Conf. on Electrical Engineering and Informatics* **1** pp 130-34
[16] Yudong Zhang, Shuihua Wang, Genlin Ji and Preetha Phillips 2014 Fruit classification using Computer vision and feedforward neural network *Journal of Food Engineering* pp 43-67