International Journal of Computer (IJC)
ISSN 2307-4523 (Print & Online)

http://ijcjournal.org/

CNN Transfer Learning for Automatic Fruit Recognition for Future Class of Fruit

Israr Hussain[a*], Shunquan Tan[b], Wajid Ali[c], Amjad Ali[d]

[a,b]*College of Electronic and Information Engineering, Shenzhen University, Shenzhen China*
[c]*Department of Mathematical Sciences Karakoram International University Gilgit Pakistan*
[d]*Department of Agriculture and Food Technology Karakoram International University Gilgit Pakistan*
[a]*Email: israrhussain46@gmail.com* , [b]*Email: tansq@szu.edu.cn*， [c]*Email: wajid.ali@kiu.edu.pk*， [d]*Email: dr.amjadali@kiu.edu.pk*

Abstract

Deep fruit recognition model learned on big dataset outperform fruit recognition task on difficult unconstrained fruit dataset. But in practice, we are often lack of resources to learn such a complex model, or we only have very limited training samples for a specific fruit recognition task. In this study we address the problem of adding new classes to an existing deep convolutional neural network framework. We extended our prior work for automatic fruit recognition by applying transfer learning techniques to adding new classes to existing model which was trained for 15 different kind of fruits. Pre-trained model was previously trained on a large-scale dataset of 44406 images. To add new class of fruit in our pre-trained model, we need to train a new classifier which will be trained for scratch, on the top of pre-trained model so, that we can re- purpose the feature learned previously for the dataset. Transfer learning using our pre-trained model has been demonstrated to give the best classification accuracy of 95.00%. The experimental results demonstrate that our proposed CNN framework is superior to the previous state-of-the- art networks.

Keywords: Deep Learning; Fruit Recognition; Transfer Learning; CNN.

1. Introduction

Files should be Pakistan is an agriculture country. Agriculture is considered the backbone of Pakistan's economy.

----

* Corresponding author.

Different types of fruits and vegetables are producing in Pakistan. In Pakistan the supermarkets and fruit shops are always very crowded because the cashiers still use the manual systems to determine the prices of the fruits. So, we have to wait for a long time to buy fruits in fruit shops and supermarkets. Fruit classification is a difficult and important task in supermarkets and fruit shops, it is necessary for the cashier to know the categories of particular fruit in order to determine its price. The use of bar-codes and issue the cashier an inventory with pictures and codes however, flipping over the inventory booklet and barcode can be overall time-consuming Rocha and his colleagues [1]. Scholars have proposed several effective solutions based on computer vision and machine learning to address similar problem during the last decade. Baltazare and his colleagues [2] first applied data fusion to nondestructive image of fresh intact tomatoes, followed by three-class Bayesian classifier. Convolutional neural networks and deep learning currently provide the best solutions to many problems in image recognition such as face recognition, fruit and vegetable recognition. Recently deep leaning-based algorithms have received significant attention because of their promising applications in various fields particularly, in object recognition and classification [3]. Some of the prior studies use deep convolutional neural network (DCNN) for fruit recognition tasks [5,6,7,8]. In this study we propose a training scheme for adding new classes to an existing framework where the dataset that was used to train the network may not be accessible. As per our knowledge, previous studies in related transfer learning problems have been constricted to the context of computer vision in numerous images understanding tasks.This study is step forward in this direction and we utilized transfer learning techniques in our prior framework [3,4] for automatic fruit recognition to add future fruit class in our pre-trained model. As per our knowledge, previous studies in related learning problems have been restricted to the context of computer vision in various image understanding tasks. Here, however, we apply the proposed method to the fruit recognition task.

2. Transfer Learning

It is difficult to collect thousands of training data samples needed to develop a deep learning-based model that will achieve acquire accuracy in practical applications. To this end, transfer learning is conceived an effective way in training custom deep learning frameworks [12]. In machine learning transfer learning is a method where, a model developed for a particular task is reused as the starting point for another task. When new dataset is small and similar to the original dataset. Since the dataset is small, it not a good idea to fine-tune the CNN due to overfitting problems. If our dataset is alike to the original dataset, we expect that the high -level features in convolution neural network to be relevant to this dataset as well. So, the best idea is to train a linear classifier on the CNN codes. If the new dataset is large enough and similar to the original dataset. We have sufficient training dataset so; we are more confident that we won't overfit the model if we will try to fine-tune through the full network. New dataset is small but very different from the original dataset. Since the dataset is small it best to train linear classifier. In case if the dataset is very different from the original dataset. It might not be a good option to train the classifier form the top of the network which contains more dataset-specific features. Instead, it might work better to train SVM classifier from activations somewhere earlier in the network. When the new dataset is large and different from the original dataset. Since the dataset is very large, we may expect that we can afford to train a CNN from scratch. However, in practice it is very useful to initialize with weights from a pre-trained model. In this case we have enough dataset and confident to transfer learning through the entire network. Some of the prior deep learning-based approaches [9,10] have been used transfer learning techniques for fruit recognition task to reduce the parameters and cost of computation in training process. Target dataset is small and similar to the base training dataset

3. Pre-Trained CNN Network Architecture

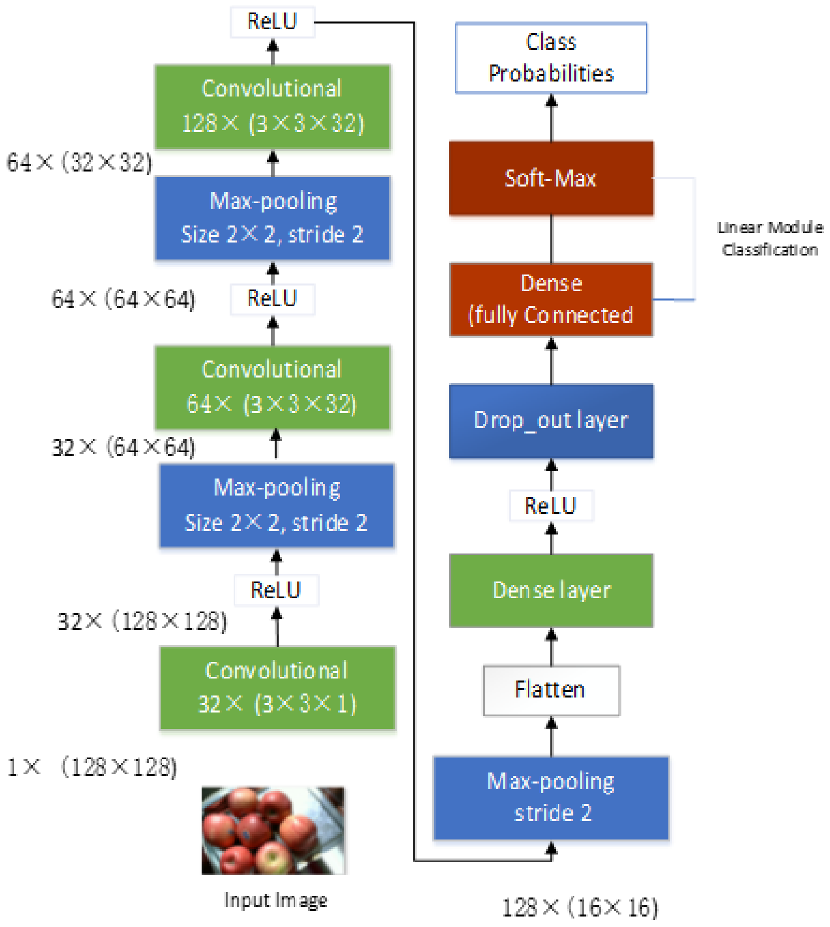

Figure1: pre-trained network architecture

Our pre-trained model is a two-track deep neural network framework. The first track consists of deep convolution neural network with max-pooling to enhance system ability, whereas the second track comprised of fully connected layer. The network has four layers of hidden neurons (three convolutional-pooling and one fully connected), apart from a final dense layer of output neurons (the input is not considered as a layer). The input contains 150 × 50 × 3 neurons, representing the RGB value for a 150 × 50 × 3 image. The first convolution-pooling layer uses a local receptive field (convolutional kernel) of size 3 × 3 with a stride length of 1 pixel to extract 32 feature maps, followed by a max. pooling operation conducted in a 2 × 2 region, the second and third convolution-pooling layers use the same 3 × 3 local receptive field (kernel) resulting 64 and 128 features maps respectively whereas another parameter remain unchanged. The fourth layer is fully-connected layer with 64 ReLU (rectifier linear unit) neurons and the output layers have 15 SoftMax neurons that corresponding to the 15-various categories of fruits. The three convolutional-pooling layers also use ReLU activation functions. Four main operations are performed in the Conv.Net as shown in Figure.1.

4. Proposed transfer learning approach

In this study we propose a training strategy for adding new classes to an existing network where the dataset that was used to train re-trained the output dense layer only the network cannot be accessed.

4.1. Re-trained the output dense layer only

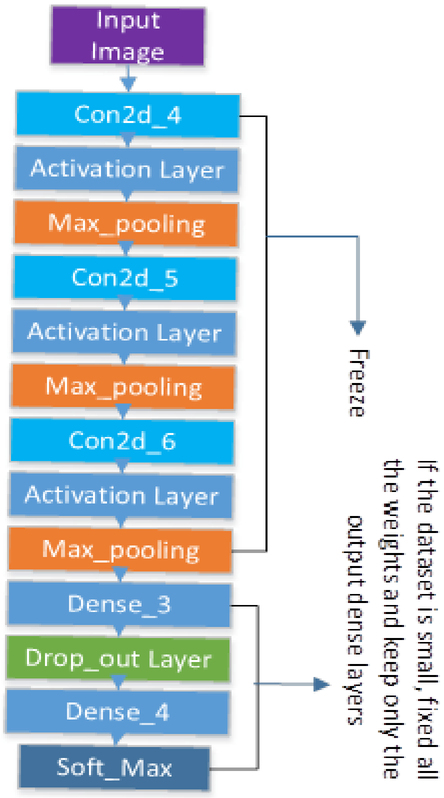

Figure 2: Transfer learning approach-1 with pre-trained model

Freeze all the layers in the network except the last two layers (weights are non-trainable)

for layer in base_model. layers [:2]:

new_model.add (layer)

Freeze earlier Conv layers earlier in the network to make sure that any previous robust feature learned by our pre-trained model are not destroyed. Then start the fine tuning by only allow the gradient to back propagate through only the fully connected layers. This technique helps us our network to boost up. Here we used CNN as a feature extractor. Then we use these features and send them to dense layers which are trained according to our dataset. The output layer is replaced with a SoftMax layer relevant to our problem. The output layer in a CNN is a SoftMax layer for 15 categories of fruits. We just trained the weights of dense layers and try to identify the new class of the fruit. Now add a dense layer at the end and we will only train this dense layer.

4.2. Freeze the weights of first few layers

In this case we optionally unfreeze some of the convolutional layers in the network and perform the second phase of training. The initial layers of convolutional neural networks just learn the general features like edges, dots, curves Deeper part extract special features. So, we can freeze first few layers. In our case we freeze the weights of first 5 layers of the CNN network Where we retrain the subsequent layers. This is because the first few layers capture universal features like shape, edges, color that are also relevant to our new problem.

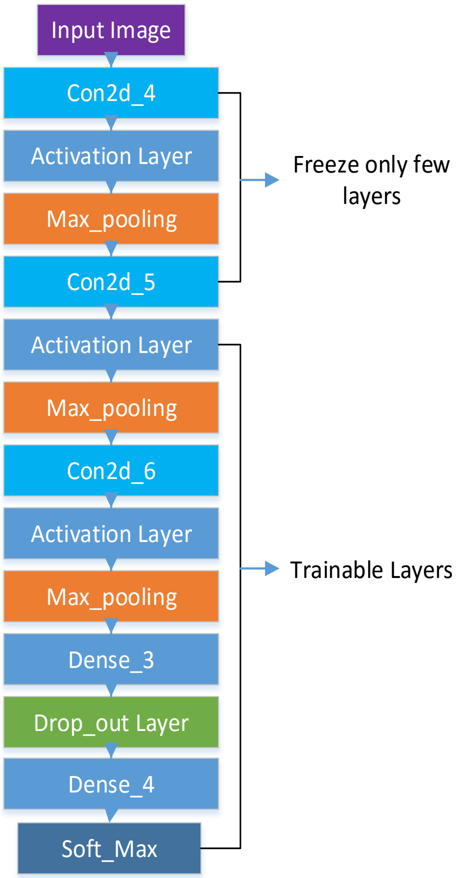

Figure 3: Transfer learning approach-2 with pre-trained model

To set the first 5 layers of the network to non-trainable (weights will not be uploaded).

```
for layer in base_model. layers [:5]
    layer. trainable=False
```

5. Experimental process and result analysis

5.1. Dataset Collection Considerations

Dataset used in our pre-trained model was comprising of 44406 images for 15 different categories of fruits [9]. While the new class of cherry fruit consisting of 450 images. The images were collected at different day times of the day and in different days for the same category. These features increase the dataset variability and represent more realistic scenario. The Images had large variation in quality and lighting. Illumination is one of those variations in imagery. In fact, illumination can make two images of same fruit less similar than two images of different kind of fruits. We were used our own intelligent weight machine and camera to captured all images. The fruit dataset was collected under relatively unconstrained conditions. There are also images with the room light on and room lights off, moved the camera and intelligent weight machine near to the windows of our lab than open windows, closed windows, open window curtains, closed curtains. For a real application in a supermarket, it might be necessary to cope with illumination variation, camera capturing artifacts, specular reflection shading and shadows. Below are the few conditions which we considered during collecting data samples.

- Pose Variations with different categories of fruits
- Variability on the number of elements of fruits
- Used HD camera with 5-megapixel snapshots
- Same color but different Category fruits images with illumination variation
- Cropping and partial occlusion
- Different color same category fruit images
- Different lighting conditions it can easily affect the recognition system
- Six different kind of apple fruit images
- Three categories of Kiwi fruit images
- Partial occlusion with hand

The network was trained with the stochastic gradient descent (SGD) algorithm with learning rate of 0.0001, momentum 0.9. Weights in each convolutional layer and fully connected layers are initialized random numbers generated from zero mean Gaussian distribution with standard deviation of 0.01. The learning rate are initialized to 0.0001 and set to decrease 10% for each 5,000 iterations. The momentum and weights decay are 0.9.

5.2. Software and hardware

Software and Hardware: Tensor-flow platform was selected to implement the proposed CNN framework which was configured to use GPU Quadro K4200 for training. A mini-batch of 32 were used as input for each iteration.

Table 1: Hardware and software configuration

GPU Model	Quadro K4200/PCIe/SSE2
Processor	Inter Xeon® CPU E5-2609v3 @1.90GHz×6
Memory	15.6 GiB
Operating system	Window10
Disk	587.2
Programming Language	Python 3.6/OpenCV

5.3. Experiments results with transfer learning

To add a new class of fruit in our pre-trained model for 15 categories of fruits We set the first 5 layers to non-trainable means the weights will not be updated during train the model for new class of fruit. Our experimental results illustrate the behavior of our proposed framework in term of accuracy and loss curves. The accuracy and loss curves show that the propose approach is capable of automatically recognizing the cherry fruit with a high accuracy of 95%. In our experiments we set the first 5 layers to non-trainable means the weights will not be updated during train the model for new class of fruit. The first few layers of convolutional neural networks extract the very basic features from the input such as edges, curves, dots and line while the deeper part extract more complex features. In our experiments after freeze first, few layers when we trained the model for new class of cherry fruit, we have got an accuracy of 95% as shown in Figure. 4 and Figure. 5. The training and validation accuracy curves in figure. 4 illustrated that the proposed model accuracy could be improved further with increasing the number of training epochs. We also have evaluated the proposed network performance on other kind of fruit categories with our transfer learning techniques and achieved an excellent detection performance. As shown in below validation and training accuracy curves for 25 epochs. We kept same all the other hyper-parameters same as for our pre-trained model for 15 different kind of fruits. We just freeze some initial layers and replaced the output SoftMax layer. As shown in validation and training curves the validation and test accuracy overlapping each other with the increasing number of epochs. This means the model has generalized very well. If the validation and training curves go away from each other its mean the model is overfitting and need to apply overfitting techniques to overcome overfitting problem. As shown in figure. 4, the first test accuracy was greatly elevated to a level of 95% with 25 epochs.

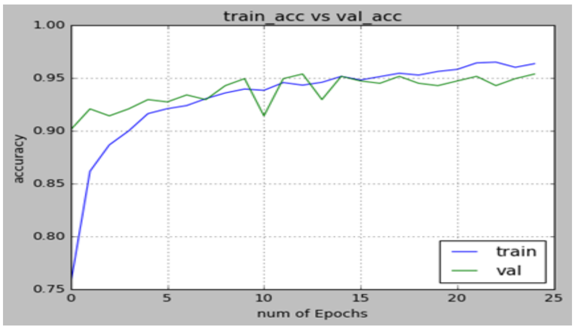

Figure 4: Validation and training accuracy curve

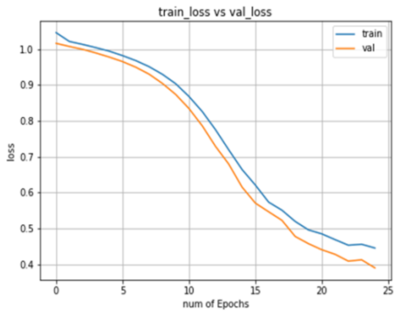

Figure 5: Validation and training loss curves

5.4. Classification report table

Table 2: Classification report table [4]

Label	Fruit Name	Precision	Recall	F1-score	Support
0	Plumb	1.00	1.00	1.00	445
1	Apple	1.00	0.97	0.99	557
2	Banana	0.98	1.00	0.99	418
3	Carambola	1.00	0.99	1.00	326
4	Guava	1.00	1.00	1.00	644
5	Kiwi	1.00	1.00	1.00	845
6	Mango	0.99	0.99	0.99	575
7	Muskmelon	1.00	1.00	1.00	318
8	Orange	0.98	0.99	0.98	314
9	Peach	1.00	1.00	1.00	394
10	Pear	1.00	0.99	1.00	361
11	Persimmon	0.99	1.00	1.00	275
12	Pitaya	0.99	0.99	0.99	411
13	Pomegranate	1.00	0.99	1.00	326
14	Tomatoes	0.99	1.00	0.99	595
15	Cherry	0.95	0.95	0.95	450
Avg/total		**0.99**	**0.99**	**0.99**	**7254**

Recognition probabilities for test image and their corresponding recognition rates by our proposed framework for each fruit category are shown in Table.2. We also highlighted the detection accuracy for new class of cherry fruit in table.2 (label 15). The recall is the capability of the framework to recognize all the positive samples. The row with light blue color shows the recognition accuracy for our new class of cherry fruit by our transfer learning approach.

6. Conclusion

In this paper we extended our prior work for automatic fruit recognition based on DCNN for commercial source trace system [4] by utilizing transfer learning techniques. Our pre-trained model was developed for 15 different categories comprising of 44406 images. To add a new class of fruit in our pre-trained model for 15 categories of fruits. We used two approaches in the first approach we set the first five layers of our-pre-trained network to non-trainable means the weights will not be updated during train the model for new class of fruit. In the second approach we freeze all the convolutional layers except the fully connected layers. After freeze first few layers in our pre-trained network when we trained the model for new class of cherry category which did include in our pre-trained model and we have noted that we achieved an excellent performance of 95% detection accuracy for cherry fruit.

References

[1]. A. Rocha, D. C. Hauagge, J. Wainer, S. Goldenstein ‘Automatic fruit and vegetable classification from images’, Computers and Electronics in Agriculture, 2010, 70, pp. 96– 104.

[2]. A. Baltazar, J. I. Aranda, G. Gonzalez Aguilar, Bayesian classification of ripening stages of tomato fruit using acoustic impact and colorimeter sensor data, Computers and electronics in agriculture’, 2008, 60, pp. 113–121.

[3]. S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, S. Iyengar, ‘A survey on deep learning: Algorithms, techniques, and applications, ACM Computing Surveys (CSUR)’, 2018, 51, pp.51, 1–36

[4]. Hussain, Israr, Qianhua He and Zhuliang Chen. ‘Automatic Fruit Recognition Based on DCNN for Commercial Source Trace System. International Journal of Computer Science & Applications’, 2018, 8, pp: 01-14.

[5]. I. Hussain, W. L. Wu, H. Q. Hua, N. Hussain, ‘Intra-class recognition of fruits using dcnn for commercial trace back-system, in: Proceedings of the 4th International Conference on Multimedia Systems and Signal Processing (MSSP)’, 2019, pp. 194–199.

[6]. Mure¸San, Horea, and Mihai Oltean. ‘Fruit recognition from images using deep learning’, Acta Universitatis Sapientiae Informatica’, 2018, 10, pp. 26-42.

[7]. Wan, Shaohua, and Sotirios Goudos. ‘Faster R-CNN for multi-class fruit detection using a robotic vision system’, Computer Networks, 2010, 168, pp. 107036.

[8]. Ludwig, I. and Ludwig A. W. W ‘Kondo effect induced by a magnetic field’, Phys. Rev. B, 2001, 64, pp. 045328

[9]. Israr Hussain, Qianhua He, Zhuliang Chen, & Wei Xie. (2018). Fruit Recognition dataset (Version V 1.0) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.1310165

[10]. Alvarez-Canchila, O. I., D. E. Arroyo-Pérez, A. Patino-Saucedo, H. Rostro González, and A. Patino-Vanegas. ‘Colombian fruit and vegetables recognition using convolutional neural networks and transfer learning’, In Journal of Physics: Conference Series, 2020., vol. 1547, no. 1, p. 012020. IOP Publishing,

[11]. Siddiqi, Raheel. ‘Effectiveness of Transfer Learning and Fine Tuning in Automated Fruit Image Classification’, In Proceedings of the 2019 3rd International Conference on Deep Learning Technologies, 2019 pp. 91-100.

[12]. Tan, Chuanqi, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. ‘A survey on deep transfer learning’, In International conference on artificial neural networks, 2018, pp. 270-279.