**FruitTL: A Transfer Learning Approach towards Classification of Inidan Fruits**

*Ankan Bhattacharyya1, Shibaprasad Sen2,\**

*Affiliations:*

*1, University of Kentucky, USA*

*2, Techno India Main Campus, Salt Lake, India*

*\*Corresponding author*

*Emails:*

*Ankan Bhattacharyya:* [*abh242@uky.edu*](mailto:abh242@uky.edu)

*Shibaprasad Sen:* [*shibubiet@gmail.com*](mailto:shibubiet@gmail.com)

*Abstract: The manuscript discuss a transfer learning approach to classify fruits. A pretrained VGG16 model has been applied with transfer learning approach to the FruitNet dataset. The model has obtained a training accuracy of 99%, which beats the state of the art models that has been tested on the FruitNet dataset. The code is available in* [*GitHub*](https://github.com/iamankan/fruits-ml/blob/conference1/FruitmlTL.ipynb)*.*

*Keywords: Transfer learning, VGG16, CNN*

1. Introduction

The turn-over for the fruit industry with respect to the agricultural output is good [15-17]. So, faster and more efficient methods for classifying fruits are needed. Machine learning is one of the way, as discussed later in this manuscript of how machine learning/deep learning has become an integral part of the industry. Agriculture is also a major source of employment in some areas of the world. Still there are hand-picking methods for classifying fruit quality and size. This requires a lot of human effort and the amount that are being paid to the farmers and the other people in the chain, is not worth it. Lack of knowledge for fertilizers, soil maintenance and other factors can affect the crop quality. In the demand supply chain, the requirement of the food for consumers are increasing at a faster rate than that of the production. One of the key factors in production of food is the quality of the food. That is why machine learning techniques, which are unbiased and faster are required to classify the food in different aspects like, quality, size, disease, etc.

Agriculture activities usually covers three spectrum, namely, pre-harvesting, harvesting and post harvesting [17].

Data in the fruits and vegetable domains are mainly in form of photos taken with cameras that are easily maneuverable. So, digital image processing is required to classify and detect the quality and other things, like, texture, color, etc. Erdenee et. al. [20] has done identification of land using image processing. Tiwari et. al. [21] has evaluated nitrogen in plants. Krishna et. al. [22] has detected pest infected areas. Patil et. al. [22] detected plant disease using properties like texture, shape and color. Computer vision is one of the best ways to do these kind of tasks because they mimic the human vision very closely [24-26]. Bhargava et. al. [15] explains very good how CT scan has been used to develop image datasets for fruits and vegetables.

With the advent of the data and computational power, today deep learning is being used in different practice, starting from handwriting recognition [2, 4] to generating images from text, like DALL-E [8]. Turing Award Winners Yan LeCun, Yoshua Bengio, and Geoffery Hinton, have developed DNN (Deep Neural Network) model which is denser than ANN (Artificial Neural Network) (9) and is well capable to extract complex features at an abstract level. With the advent of DNN, many models have been created over time and the most popular model is the CNN. As a result, many researchers have started experimentation with CNN for the solution of various complex pat- tern recognition problems. Ciresan et al. have experimented to improve the overall classification accuracy of online handwritten characters by using CNN [10]. Sen et al. have shown the usefulness of the CNN model towards online handwritten Bangla character recognition [11]. Authors in [12] have shown the procedure to recognize English characters and digits by using multi-CNN. Bal- dominos et al. [13] have used evolutional CNN as an application of handwriting recognition. Mehrotra et al. have introduced an offline strategy to recognize online handwritten Devanagari characters using CNN [14]. Bhattacharyya et. al. [19] have used deep learning to recognize human facial expression from infrared images of human face.

The present work classifies the fruits in India using Transfer Learning of VGG-16. The paper has been written as follows: Section 2 explains the dataset. Section 3 explains the methodology and the description of the model used. Section 4 discusses the results and Section 5 concludes with a future scope of the present work, as an extension.

1. Dataset

The dataset [1] has been collected using high resolution mobile phone rear camera. Meshram et. al. [1] has not mentioned anything about the specifications of the mobile camera. However, the specification of the images taken has been mentioned to be .jpg images of dimension 3024x3024. The images has been resized down to 256x256. The dataset has three subfolders, namely, Good Quality, Bad Quality and Mixed Quality fruits. For each of these categories, there are mainly 6 fruit types, namely, Apple, Guava, Banana, Lime, Orange and Pomegranate. Meshram et. al. claims that the data were taken at different lighting conditions in different background.

This manuscript deals with the fruit classifications, so, the quality of the fruits are not in the scope of the current experiment. We rearranged the data into six classes of different fruit types. The data is uploaded to the [cloud](https://drive.google.com/drive/folders/11V65XEXsYM2pZV2lERbK0ESxSARWxCMS?usp=sharing). We are experimenting with a total of 18452 data samples.

1. Methodology

We have made a split of the data into training and testing samples. The ratio of the training to the testing samples is 8:2. Further the training data is split into 90% training and 10% validation set. All the splitting done for this experiment are random and shuffled, so that the distribution is uniform. So, after the splitting, the following are the data counts:

|  |  |
| --- | --- |
| Training | 11629 |
| Testing | 5534 |
| Validation | 1289 |

We have used VGG16 as the pre-trained deep learning model. Chakraborty et. al. [2] showed that using pre-trained weighted neural network is better that running a CNN from scratch, for a small scale of data. For our experiment, the scale of data is medium and cannot be considered as large.

A high level workflow diagram is shown in Fig. 1.

|  |
| --- |
| Diagram  Description automatically generated |
| Fig. 1. High level workflow diagram |

So, we approached the problem with transfer learning. A brief description of the VGG-16 is given in the following paragraph.

VGG-16 is a widely used CNN model. The network has 13 convolution layers and 3 dense fully connected layers. Hence it is called VGG-16. Fig. 2 illustrates the high-level block diagram for VGG-16. The convolution layers have 3 x 3 filter. Batch normalization has been used in this model. This is for achieving stability without overestimation or underestimation. The output layer has 1000 nodes. This network, when considered for transfer learning is assumed to have been pretrained with weights on imagenet dataset [3].

|  |
| --- |
| Diagram  Description automatically generated |
| Fig. 2. A high level block diagram for VGG-16 |

Transfer learning is the process where the pre-trained network is used with pretrained weights. As seen in Fig. 2, the last stack of convolution layer produces very complex features. We use this feature, flatten them and pass it through 2 dense layer. The first layer contains 1024 neurons and the final layer contains 6 neurons, which represent 6 classes of the fruits in the dataset. Since the model is already pretrained, we don’t retrain the layers of VGG-16 and hence freeze those layers. The only thing that is trained is the two dense layers that we added to the last entries of the stack. A high level diagram is illustrated in Fig. 3. We term this model as FruitTL, which is an acronym for Transfer Learning on FruitNet dataset [1].

|  |
| --- |
| Diagram  Description automatically generated |
| Fig. 3. A high level block diagram for transfer learning using pretrained VGG-16 |

The training of the model is done over 20 epochs, in Google Colab. The dense layer of 1024 neurons uses the Rectified Linear Unit (ReLU) activation function [5]. The ReLU function is illustrated in Fig. 4.

|  |
| --- |
| A Practical Guide to ReLU. Start using and understanding ReLU… | by Danqing  Liu | Medium |
| Fig. 4. Rectified Linear Unit activation function |

For the last classification dense layer, having 6 neurons, Softmax activation function [5] has been used. The Softmax function is illustrated in Fig. 5.

|  |
| --- |
| SoftMax Activation Function-InsideAIML |
| Fig. 5. Softmax activation function |

The model saved the weights in the overall training process at its best performance. Quantitatively, the best performance is considered as the minimum loss value. The loss function used in the training process is categorical cross-entropy [6]. Categorical cross entropy is used because there are more than two data (hence not binary). Th optimizer used for the training is the Adam optimizer [7].

1. Results and discussions

We performed VGG-16 transfer learning approach as discussed in the earlier sections. We have calculated categorical accuracy with two optimizers, one with Adam and the other with RMSProp. This is shown in Table. 1.

|  |  |  |
| --- | --- | --- |
| Optimizer | Categorical Accuracy (%) | |
| Training | Validation |
| Adam | 99.89 | 98.99 |
| RMSProp | **99.99** | 98.84 |
| Table 1. Result of Adam and RMSProp optimizer with pre-trained VGG-16. | | |

Clearly, RMSProp produces the best Training result. The training graph for RMSProp and Adam is shown in Fig. 6.

|  |  |
| --- | --- |
| Chart, line chart  Description automatically generated | Chart, line chart  Description automatically generated |
|  | Chart, line chart  Description automatically generated |
| (a) | (b) |
| Fig. 6. Training and loss accuracy graph with (a) Adam optimizer and (b) RMSProp optimizer | |

The confusion matrix on 5534 test images are shown in Fig 7.

|  |
| --- |
|  |
| Fig. 7. Confusion matrix for the FruitTL on test data |

The other measurement matrix are illustrated in Fig. 8.

|  |
| --- |
| A picture containing text, receipt  Description automatically generated |
| Fig. 8. Other measurement matrix during training. |

Some other results for comparison are given in Table 2.

|  |  |  |
| --- | --- | --- |
| Author | Remarks | Reported Accuracy (%) |
| Chen et. al. [27] | Fruit classification using multi-optimization  CNN. | 99% |
| Adeel et. al. [28] | Deep features selection framework for grape leaf diseases recognition | 99% |
| Kaggle [29] | Kaggle code on FruitNet dataset | 98.87% |
| Kaggle [30] | Kaggle code on FruitNet dataset | 99.39% |
| Our work | VGG-16 transfer learning on FruitNet dataset | 99.99% |

1. Conclusion and future scope

To conclude, the Transfer learning approach works very well on dataset of decent size. The images need to be preprocessed more. If camera specifications were mentioned, then the neural networks could have been tailored according to the specifications. As a future scope, we would like to make the datset into detecting whether the quality of the fruits are good or bad. This will help us to build a two stage robust pipeline. There are works that exists with a sig\ngle pipeline. Those work lack in the tight coupling of training. This is because those are trained won bad and good quality of specific fruits like, apple\_good and apple\_bad. Unlike those, we are planning to build a two stage pipeline where the first stage detects the fruits and the final stage takes the classified fruit and classify them as bad or good. This will make the system robust and can scale upto any number of fruits and any type of quality, which is tailored towards fine-grained classification of the quality.

1. References

[1] Vishal Meshram, Kailas Patil, FruitNet: Indian fruits image dataset with quality for machine learning applications Data In Brief, Volume 40 (2021), [10.1016/j.dib.2021.107686](https://doi.org/10.1016/j.dib.2021.107686)

[2] Chakraborty, R., Saha, S., Bhattacharyya, A., Sen, S., Sarkar, R., Roy, K. (2021). Recognition of Online Handwritten Bangla and Devanagari Basic Characters: A Transfer Learning Approach. In: Singh, S.K., Roy, P., Raman, B., Nagabhushan, P. (eds) Computer Vision and Image Processing. CVIP 2020. Communications in Computer and Information Science, vol 1377. Springer, Singapore. https://doi.org/10.1007/978-981-16-1092-9\_45

[3] <https://www.image-net.org/>

[4] Bhattacharyya, A., Chakraborty, R., Saha, S. et al. A Two-Stage Deep Feature Selection Method for Online Handwritten Bangla and Devanagari Basic Character Recognition. SN COMPUT. SCI. 3, 260 (2022). https://doi.org/10.1007/s42979-022-01157-2

[5] <https://github.com/keras-team/keras/blob/v2.9.0/keras/activations.py#L273>

[6] <https://www.tensorflow.org/api_docs/python/tf/keras/metrics/categorical_crossentropy>

[7] <https://arxiv.org/abs/1412.6980>

[8] <https://doi.org/10.48550/arXiv.2102.1209>

[9] Krizhevsky, A., Sutskever, I., Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks. Communications of the ACM 60(6), 84-90 (2012).

[10] Cireşan, D. C., Meier, U., Gambardella, L. M., Schmidhuber, J.: Convolutional neural network committees for handwritten character classification. In: Proceedings of the International Conference on Document Analysis and Recognition, pp. 1135–1139 (2011).

[11] Sen, S., Shaoo, D., Paul, S., Sarkar, R., Roy, K.: Online Handwritten Bangla Character Recognition Using CNN: A Deep Learning Approach. In: Advances in Intelligent Systems and Computing, pp. 413-420 (2018).

[12] Pham, D. V.: Online handwriting recognition using multi convolution neural networks. Lecture Notes on Computer Science, LNCS, vol. 7673, pp. 310–319 (2012).

[13] Baldominos, A., Saez, Y., Isasi, P.: Evolutionary convolutional neural net- works: An application to handwriting recognition. International Journal of Neurocomputing 283, 38–52 (2018).

[14] K. Mehrotra, S. Jetley, A. Deshmukh, S. Belhe, ”Unconstrained handwritten Devanagari character recognition using convolutional neural networks,” in: Proceedings of the 4th International Workshop on Multilingual OCR, pp. 1-5 (2013).

[15] A. Bhargava, A. Bansal Fruits and vegetables quality evaluation using computer vision: a review J. King Saud Univ. Comput. Inf. Sci., 33 (2021), pp. 243 257, [10.1016/j.jksuci.2018.06.002](https://doi.org/10.1016/j.jksuci.2018.06.002)

[16] S. Behera, A. Rath, A. Mahapatra, P. Sethy Identification, classification & grading of fruits using machine learning & computer intelligence: a review J. Ambient Intell. Humaniz. Comput. (2020), [10.1007/s12652-020-01865-8](https://doi.org/10.1007/s12652-020-01865-8)

[17] V. Meshram, K. Patil, V. Meshram, D. Hanchate, S. Ramteke Machine learning in agriculture domain: a state-of-art survey Artif. Intell. Life Sci. (2021), [10.1016/j.ailsci.2021.100010](https://doi.org/10.1016/j.ailsci.2021.100010)

[18] A deep learning model for classifying human facial expressions from infrared thermal images

[19] Bhattacharyya, A., Chatterjee, S., Sen, S. et al. A deep learning model for classifying human facial expressions from infrared thermal images. Sci Rep 11, 20696 (2021). https://doi.org/10.1038/s41598-021-99998-z

[20] Erdenee, B., Ryutaro, T., Tana, G., 2010, Particular Agricultural Land Cover Classification Case Study Of Tsagaannuur, Mongolia. In: IEEE International Geoscience & Remote Sensing Symposium, 3194-3197.

[21] V.K. Tewari, A.K. Arudra, S.P. Kumar, V. Pandey, N.S. Chandel Estimation of plant nitrogen content using digital image processing Int. Commission Agricu. Biosyst. Eng., 15 (2) (2013), pp. 78-86

[22] M. Krishna, G. Jabert Pest control in agriculture plantation using image processing IOSR J. Electron. Commun. Eng. (IOSR-JECE), 6 (4) (2013), pp. 68-74

[23] J.K. Patil, R. Kumar Advances in image processing for detection of plant diseases J. Adv. Bioinf. Appl. Res. ISSN, 2 (2) (2011), pp. 135-141

[24] S. Naik, B. Patel A machine vision based fruit classification and grading: a review Int. J. Comput. Appl., 170 (9) (2017), pp. 22-34

[25] S.R. Dubey, A.S. Jalal Application of image processing in fruits and vegetables analysis: a review J. Intell. Syst., 24 (4) (2015), pp. 405-424

[26] B. Zhang, Z. Huang, J. Li, C. Zhao, S. Fan, J. Wu, C. Liu Principle, developments and applications of computer vision for external quality inspection of fruits and vegetables: a review Food Res. Int. (2014), pp. 326-343

[27] Chen, X., Zhou, G., Chen, A. et al. The fruit classification algorithm based on the multi optimization convolutional neural network. Multimed Tools Appl 80, 11313–11330 (2021). https://doi.org/10.1007/s11042-020-10406-6

[28] Adeel, A, Khan, MA, Akram, T, et al. Entropy-controlled deep features selection framework for grape leaf diseases recognition. Expert Systems. 2020; 1– 17. <https://doi.org/10.1111/exsy.12569>

[29] <https://www.kaggle.com/code/gpiosenka/indian-fruit-18-classes-f1-score-99>

[30] <https://www.kaggle.com/code/tarushfx/fruit-quality>