Fruit Recognition from Images Using Machine Learning

M.G.T.M. Kumara (Leader)	AS2017418
J.M.D.N.L. Jayamaha	AS2017391
J.D.K.N. Jayakodi	AS2017389
A.H.R. Lankapriya	AS2017423
T.D.K. Buddika	AS2017332

Introduction

Sri Lanka is one of the most renown sort of tourist destinations in the world. The diversity of temperature, vegetation and scenery grab a large sense towards Sri Lanka and becomes one of the demanding countries among nature lovers and fans of wildlife because tourism offers safaris, nature walks and camping and that would be a marvelous experience for them. While enjoying, they prefer to taste local food specially fruits as here in Sri Lanka, we have difference varieties and types of fruits which are tasty and have medicinal value. But we never really give them the fact that how our island is home for an immense variety of fruits until many of the tourists asking about what fruits we have to try them while they are spending their vacations. Most get surprised and made curiosity after seeing those various fruits in the Sri Lankan markets which are very rare and some names are not even heard before. So, we are going to make an app to promote local fruits among tourists using image processing technique to give better understanding about fruits and its value just using an image of a fruit.

Significance of the research

Sri Lanka is famous for its various types of fruits because different areas of the country have its own weather conditions. Foreigners and also the local travelers basically buy fruits while there are going on a journey. But there are many types of them and specially foreigners can't remember those names and significance value of every fruit. So, if they have an app which gives them the information of fruits which they can buy in the local market, it will help them to buy the fruits which they want the most and will help to deal with local market sellers who don't know the foreign languages. As a part of research area, we develop this fruits classification using convolutional neural network and deep learning technique which is a class of machine learning.

Literature review

There are several previous attempts to use neural networks and deep learning for fruits recognition. Although many researchers have tackled the problem of fruit detection, such as the works presented in [1,2,3,4,5,6] the problem of creating a fast and reliable fruit detection system persists, as found in the survey by [7]. This is due to high variation in the appearance of the fruits in field settings, including color, shape, size, texture and reflectance properties. Wang et al. [4] examined the issue of apple detection for yield prediction. They developed a system that detected apples based on their color and distinctive specular reflection pattern.

Further information, such as the average size of apples, was used to either remove erroneous detections or to split regions that could contain multiple apples. Bac et al. [5] proposed a segmentation approach for sweet peppers. They used a six band multi-spectral camera and used a range of features, including the raw multispectral data, normalized difference indices, as well as entropy-based texture features.

Experiments in a highly controlled glasshouse environment showed that this approach produced reasonably accurate segmentation results. However, the authors noted that it was not accurate enough to build a reliable obstacle map.

Hung et al. [6] proposed the use of conditional random fields for almond segmentation. They proposed a five-class segmentation approach, which learned features using a Sparse Auto Encoder (SAE).

These features were then used within a CRF framework and was shown to outperform previous work. They achieved impressive segmentation performance, but did not perform object detection.

Furthermore, they noted that occlusion presented a major challenge. Intuitively, such an approach is only able to cope with low levels of occlusion.

Yamamoto et al. [3] performed tomato detection by first performing color-based segmentation. Then, color and shape features were used to train a Classifier and Regression Trees (CART) classifier.

produced a segmentation map and grouped connected pixels into regions. Each region was declared to be a detection and to reduce the number of false alarms. They trained a non-fruit classifier using a random forest in controlled glasshouse environments.

Recently, deep neural networks have made considerable progress in object classification and detection [8,9,10]. The state-of-the-art detection framework on PASCAL-VOC [11] consists of two stages.

The first stage of the pipeline applies a region proposal method, such as selective search [12] and edge box [13] to extract regions of interest from an image and then feed them to a deep neural network for classification.

In real outdoor, a single sensor modality can rarely provide the needed information to detect the target fruits under a wide range of variations in illumination, partial occlusions and different appearances.

This work follows the same approach and demonstrates the use of convolutional neural network (CNN) based fruit detection system and how it outperforms, as we show in the following sections.

Research question

Aim

- Fruits recognition algorithm was developed based on convolution neural network (CNN).
- Providing an efficient, accurate application to identify the fruits of Sri Lanka for both local and foreign tourists.

Objective

- Development of a machine learning algorithm to identify the few varieties of fruits in Sri Lanka.
- Provide a short description of each fruits to user.
- Training of high accuracy and high-performance machine learning algorithm to image detection.

Methodology

Describe how data will be collected

Here we are listed giving Sri Lankan endemic fruits as priority and the images were obtained by filming the fruits while they are rotated by a motor and then extracting frames. Fruits were planted in the shaft of a low speed motor (3rpm) and a short movie of 20 seconds was recorded. Behind the fruits we placed a white sheet of paper . However due to the variations in the lighting conditions, the background was not uniform and we wrote a dedicated algorithm which extract the food from back ground. This algorithm is of flood fill type; flood fill is an algorithm that will determine the area connected to given starting pixel in an image such that all pixels included in this set share the same color as that starting pixel.it is used to color an entire area of connected pixel with same color. Like that we extracted fruits from white background. The resulted dataset has approximately 7800 images of fruits spread across more than 70 labels.

Other than that, here we also used Fruits-360 dataset. Fruits-360 has 69905 images of fruits spread across more than 101 labels. The dataset which is publicly available on GitHub and Kaggle.

Describe what machine learning methods will be used in a project

Under supervised learning method Machine learning techniques and select a framework, such as Keras and TensorFlow, to train CNN and track its progress. At this time problem perform with large dataset, so in this stage deep learning algorithm requires instead of machine learning algorithms.

Deep learning is a class of machine learning algorithms that use multiple layers that contain nonlinear processing units. Each layer uses the output from the previous layer as input. Deep learning algorithms use more layers than shallow learning algorithms. Convolutional neural networks are classified as a deep learning algorithm. These networks are composed of multiple convolution layers with aa few fully connected layers. also, they make use of pooling. In area of image recognition and classification, we have chosen deep neural networks.

Convolutional neural networks (CNN)

CNN is comprising various convolutional and pooling layers that resembles human visual systems. Convolutional neural networks are part of the deep learning algorithms, also deep learning is a main branch of machine learning algorithms, such a network can be composed of convolutional layers, pooling layers, ReLU layer, layer then a Pooling layer then one or more convolutional layer and then finally one or more fully connected layer, A characteristic that sets apart the CNN from a regular neural network is taking into account the structure of images while processing them, regular network converts the input in a one-dimensional array which makes the trained classifier.

TensorFlow library

TensorFlow as backend. For model training and testing the network.this is an open source framework for machine learning for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays called tensors.

Keras

Similar as TensorFlow. Both of Keras and TensorFlow can use as numerical computation. It is capable of running on top of TensorFlow, It easy to develop deep learning models regardless of the computational backend used. Keras contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image. In addition to standard neural networks, Keras has support for convolutional and recurrent neural networks. It supports other common utility layers like dropout, batch normalization, and pooling.

Describe how your methods will be evaluated

Description of our Basic Network Architecture

In our network, colvolution layer then maps to activation layer then it maps to max pooling layer then this layer maps the fully connected layer.

Convolution Layer-convolution operation is a dot product of the original pixel values with weights defined in the filter. The results are summed up into one number that represents all the pixels the filter observed.

Activation layer - the convolution layer generates a matrix that is much smaller in size than the original image. This matrix is run through an activation layer, which introduces non-linearity to allow the network to train itself via backpropagation. The activation function is typically ReLu.

Pooling layer -Downsampleing and reducing the size of the matrix. A filter is passed over the results of the previous layer and selects one number out of each group of values.typically maximum,so called as max pooling layer

Fully connected layer - input is a one-dimensional vector representing the output of the previous layers. output is a list of probabilities for different possible labels attached to the image. label that receives the highest probability is the classification decision.

In fruits-360 dataset consists of total number of images 69905 including number of test image 31688 and also number of training 38217 images of fruits spreads more than 101 labels. All the collected images and dataset consists of 100x100 pixels images

TensorFlow 2.0, Keras 2.3.1, Matplotlib, NumPy most of libraries useful in this progress.

Gantt Chart

No		March		April				May			
		3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th
1	Project idea										
2	Project Proposal										
3	Data Collection (Gathering data set)										
4	Model Implementing and Training										
5	Validation										
6	Testing										
7	Project Report										

References

[1] Nuske S.T., Achar S., Bates T., Narasimhan S.G., Singh S. Yield Estimation in Vineyards by Visual Grape Detection;

Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS ’11); San Francisco, CA, USA. 25–30 September 2011.

[2] Nuske S., Wilshusen K., Achar S., Yoder L., Narasimhan S., Singh S. Automated visual yield estimation in vineyards. J. Field Robot. 2014;31:837–860. doi: 10.1002/rob.21541.

[3] On plant detection of intact tomato fruits using image analysis and machine learning methods.Yamamoto K, Guo W, Yoshioka Y, Ninomiya S

Sensors (Basel). 2014 Jul 9; 14(7):12191-206.

[4] Wang Q., Nuske S.T., Bergerman M., Singh S. Automated Crop Yield Estimation for Apple Orchards; Proceedings of the 13th Internation Symposium on Experimental Robotics (ISER 2012);

Québec City, QC, Canada. 17–22 June 2012.

[5] Bac C.W., Hemming J., van Henten E.J. Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper.

Comput. Electron. Agric. 2013;96:148–162. doi: 10.1016/j.compag.2013.05.004.

[6] Hung C., Nieto J., Taylor Z., Underwood J., Sukkarieh S. Orchard fruit segmentation using multi-spectral feature learning;

Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Tokyo, Japan. 3–7 November 2013; pp. 5314–5320.

[7] Kapach K., Barnea E., Mairon R., Edan Y., Ben-Shahar O. Computer vision for fruit harvesting robots-state of the art and challenges ahead. Int. J. Comput. Vis. Robot.

2012;3:4–34. doi: 10.1504/IJCVR.2012.046419.

[8] Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., et al. Imagenet large scale visual recognition challenge. Int. J. Comput.

Vis. 2015;115:211–252. doi: 10.1007/s11263-015-0816-y.

[9] Simonyan K., Zisserman A. Two-stream convolutional networks for action recognition in videos; Proceedings of the Advances in Neural Information Processing Systems; Montréal, QC, Canada.

8–13 December 2014; pp. 568–576.

[10] Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks; Proceedings of the Advances in Neural Information Processing Systems;

Tahoe City, CA, USA. 3–8 December 2012; pp. 1097–1105.

[11] Everingham M., Eslami S.M.A., van Gool L., Williams C.K.I., Winn J., Zisserman A. The pascal visual object classes challenge: A retrospective. Int. J. Comput. Vis. 2015;111:98–136.

doi: 10.1007/s11263-014-0733-5.

[12] Uijlings J.R., van de Sande K.E., Gevers T., Smeulders A.W. Selective search for object recognition. Int. J. Comput. Vis. 2013;104:154–171. doi: 10.1007/s11263-013-0620-5.

[13] Zitnick C.L., Dollár P. Computer Vision–ECCV 2014. Springer; Zurich, Switzerland: 2014. Edge boxes: Locating object proposals from edges; pp. 391–405.