AUTOMATIC FRUIT RECOGNITION BASED ON DCNN FOR COMMERCIAL SOURCE TRACE SYSTEM

Israr Hussain, Qianhua He[*], Zhuliang Chen

School of Information and Electronic Engineering
South China University of Technology Guangzhou P. R. China,510641

ABSTRACT

Automatically fruit recognition by using machine vision is considered as challenging task due to similarities between various types of fruits and external environmental changes e-g lighting. In this paper, fruit recognition algorithm based on Deep Convolution Neural Network(DCNN) is proposed. Most of the previous techniques have some limitations because they were examined and evaluated under limited dataset, furthermore they have not considered external environmental changes. Another major contribution in this paper is that we established fruit images database having 15 different categories comprising of 44406 images which were collected within a period of 6 months by keeping in view the limitations of existing dataset under different real-world conditions. Images were directly used as input to DCNN for training and recognition without extracting features, besides this DCNN learn optimal features from images through adaptation process. The final decision was totally based on a fusion of all regional classification using probability mechanism. Experimental results exhibit that the proposed approach have efficient capability of automatically recognizing the fruit with a high accuracy of 99% and it can also effectively meet real world application requirements.

KEYWORDS:

Fruit Recognition/Probability/ DCNN/ Machine vision

1. INTRODUCTION

Recognizing various kinds of fruits and vegetables is one of the most difficult task in supermarket and fruit shops, since the cashier must point out the categories of a particular fruit to determine its price. The use of barcodes has mostly overcome this problem for packed products but most of the customers want to pick their commodities by themselves instead of the packed one and thus must be weighed. This problem can be solved by issuing codes for every fruit, but here memorization is another problem which lead to error in pricing. Another alternative solution is to issue the cashier an inventory with pictures and codes, however flipping over the booklet is time consuming [1]. Some alternative solutions were proposed to overcome this problem. Veggie vision was the first supermarket produce recognition system consisting of an integrated scale and image system with a user-friendly interface [2]. Hong et al. [3] employed morphological examination to separate walnuts and hazelnuts into three groups. Baltazar at el. [4] first applied data fusion to nondestructive image of fresh intact tomatoes, followed by three-class Bayesian classifier. Pennington et al. [5] use a clustering algorithm for the classification of fruits and vegetable. Pholpho et al. [6] used visible spectroscopy for classification of non-bruised and bruised longan fruits, additionally it is combined with principal component analysis(PCA), Partial Least Square Discriminant Analysis(PLS-DA) and Soft Independent Modeling of Class Analogy (SIMCA) to developed classification models.

In recent years, Convolutional Neural Networks(CNN) have enjoyed great popularity as a means for the image classification and categorization since Krizhevsky et al. [7] won the ImageNet Large-Scale Visual Recognition Challenge(ILSVRC) 2012 competition. CNN as variant of the standard deep neural network (DNN) is characterized by a special neural network architecture consisting of alternating convolutional and pooling layers [8], it is also used to extract and combine local features from a two-dimensional input. Compared to convolutional hand-crafted feature extraction-based approaches, CNN is very effective since it is able to learn optimal features from images through adaptive process. During CNN implementation phase for image classification, researchers have to collect large-scale dataset as the ILSVERC contained more than one million images [9 ,10] for a network which however is not a trial task. One simple way to deal with this situation is to apply the CNN model that has been Pre-trained on a large-scale image data [11,12] which is called transfer performing affine transformations to the raw images. Convolutional Neural Networks (CNNs) have been established as a powerful class of models for image recognition problem. Another Fruit recognition system based on convolutional neural network approach is proposed by lei Hou in 2016[13], whereas in similar year another Fruit Detection System using deep neural network is presented in [14].

Recent advancements in deep Convolutional Neural Networks (DCNNs) have led to impressive progress in computer vision, especially in image classification. One of the most frequently used deep learning methods in image processing is the convolutional neural networks. In previous work, researchers mostly used manual method to determine the important features of each image [16] (e.g. color, size, coordinate, textures) according to their requirements. We focused on representation learning that automatically learned high level features only from users strokes in a single image instead of manually tuning existing features.

In this paper, we proposed an efficient fruit classification system using deep convolutional neural network model to extract image features and implement classification. Another contribution in this paper is that we established fruit images database comprised of 15 categories. This paper was aimed to apply Deep Convolutional Neural Network aided with the data expansion techniques to a 15 class fruit images. Images can be directly used as input to the Deep Neural network without any feature extraction. To improve the recognition performance, we proposed Deep Convolutional neural network（DCNN) model with only five-layers, it is able to learn optimal features from an large scale two dimensional input image. A series of experiments were carried out using the proposed model on our own dataset of 44406 fruit images, results showed that our classification system achieves an excellent accuracy rate of 99%.

The rest of the paper is organized as follow: Section 2 present proposed method, Section 3 Described the dataset, Section 4 comprised of experimental results, analysis, simulation and confusion matrix whereas Section 5 based on performance comparison and last section represent Conclusion.

2. PROPOSED METHOD & EXPLANATION

The methodology that involves in this scheme involves the following steps.

1. Select the fruit Images
2. Resize all fruit images from (320 ×258×3) to (150×150×3)
3. Converted all RGB images into Gray Scale
4. Transform our dataset from having shape (n, width, height) to (n, depth, width, height)

5. Split dataset into training (30082), test (6804) & validation set (7520) using keras train_test_split command
6. Convert our data type to float32 and normalize our data values from 0-255 to the range [0, 1]
7. Preprocess class labels
8. Define our model architecture
9. Compile model with SGD optimizer and categorical-cross entropy, and learning rate=0.0001
10. Fit and training data for 60 iterations and 20 iterations
11. Evaluate model on our test Dataset and calculate the confusion matrix
12. Classification has been done through Python 3.6

2.1 FLOW CHART OF FRUIT RECOGNITION SYSTEM
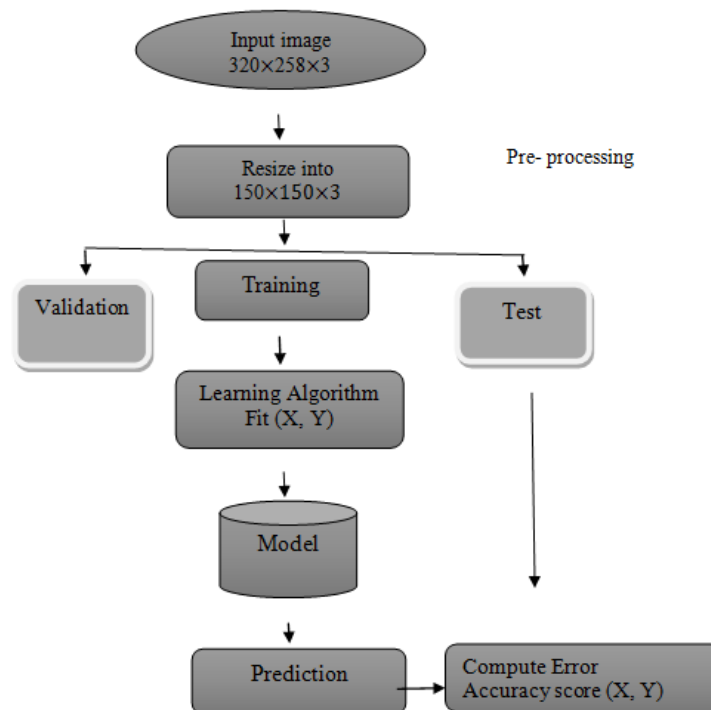
Flow Chart of fruit recognition system

2.2 PROPOSED DCNN ARCHITECTURE

In this paper we proposed a two-track deep neural network model architecture. The first track consists of deep convolution neural network with max-pooling to enhance system ability, whereas the second track comprised of fully connected layers. 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 150×150×3 neurons, representing the RGB value for a 150×150×3 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   other parameters remain unchanged. The fourth layer is

3

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

1. Convolution
2. Non-Linearity(ReLU)
3. Pooling & Sub Sampling
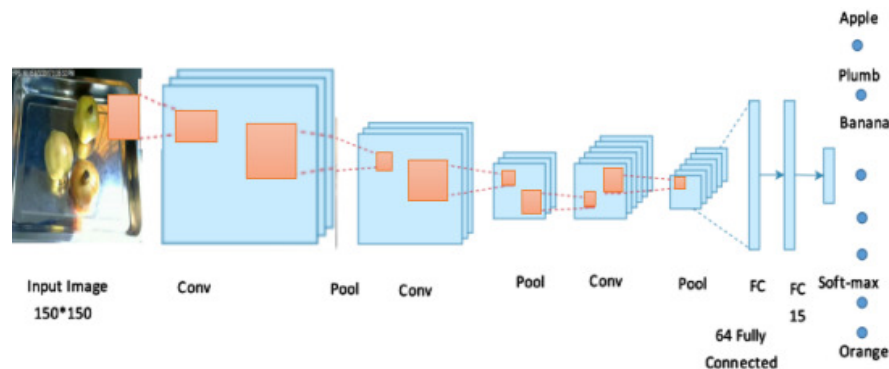4. Classification (Fully Connected Layer)

Fig.1 Proposed Deep Neural Network Architecture

Convolution in case of Conv.Net is used to extract features from the input image. Convolution preserves the spatial relation between pixels by learning image features using small squares of input image. In our model we used 3×3 filter to compute convolved features. ReLU stand for rectified linear unit is a non-linear operation. The purpose of ReLU is to introduce to non-linearity in our Conv.Net because mostly Conv.Net has to learn non-linear real-world data. In our Network, we used ReLU activation function because it trained the neural network several times faster without a significant plenty to generalization accuracy. The biggest advantage of ReLU is indeed non-saturation of its gradient, which greatly accelerates the convergence of stochastic gradient descent compared to the sigmoid / tanh functions. Another nice property is that compared to tanh / sigmoid neurons that involve expensive operations (exponentials, etc.), the ReLU can be implemented by simply thresholding a matrix of activations at zero. Fully connected layer is a traditional multi-layer perceptron that uses a SoftMax activation function in the output layer (other classifiers like SVM can be used). The convolutional and max-pooling layers extract high level features from image and used fully connected layer to classify the input images.

The network was trained with the SGD (stochastic gradient descent) algorithm with a cross-entropy cost function. As we know Adaptive optimizers are great for getting viable results quickly, but SGD + momentum is usually better for getting truly state of the art results. It just takes more effort to tune the hyperparameters properly. The dropout used to overcome the overfitting problems with dropout rates of 0.5 and 0.35 were set for the third convolutional-pooling layer and fully- connected layer respectively. The pooling layer will simply perform down sampling along with spatial dimensionality of the given input, further reducing the number of parameters within that activation. The fully-connected layers then performed the same function as found in standard ANNs and class scores are produced from the activations for classification purpose. Additionally, ReLU is used between these layers to enhance performance.

3. DATASET

Datasets plays very important role in research as it is much expensive and harder to generate because companies probably don't feel freely giving away their investment, furthermore universities don't have many resources. Therefore, our major goal was to build a dataset that contained many fruit images as feasible where each fruit is represented with many images as possible. The images were captured at the resolution of 320×258×3 pixels by using HD Logitech web camera. From experimental perspective, they were down sampled to 150×150×3 resolution in this paper. The dataset comprising of 15 different categories and the number of images for each class are listed in Table 1.

All images were stored in RGB color-space at 8 bits per channel. Images were collected at various times in different days for the same category. These features increase the dataset variability and represent more realistic scenario. The Images had large variations 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 used our own intelligent weight machine and camera to captured all images. The fruit dataset was collected under relatively unconstrained conditions. Images were also captured with different conditions i-e room light on and off, moved the camera and intelligent weight machine near to the windows of our lab, open windows, closed windows, open curtains, and closed curtains scenario. For a real application in a supermarket, it might be necessary to cope with illumination variation, camera capturing artifacts, specular reflection shading and shadows. We created all kind of real world challenges for our dataset within our lab environment. Basic few conditions which we considered during dataset collections are:

- Pose Variations with different kind of fruits
- Variable number of fruits/Vegetable
- Used HD logistic webcam to captured to images
- Same color but different Category fruits images
- Different color same category fruit images
- Different lighting conditions (i-e in fluorescent, natural light e-g within or without sunshine as some of the fruits shops and supermarkets are without sunshine so it can easily affect the recognition system.

To introduce a large amount of variations, fruit images have been captured by different fruit orientation, variability on the number of fruits, illumination, and fruit size. Different external environment variations contained in the dataset are represented in below table.1

3.1 EXISTING CATEGORIZATION CHALLENGES

Some of the existing fruit categorization challenges in our dataset are exhibit in below figures. Different fruit images may have same color, shape and sometimes it is very tough for human eye to clearly distinguish it in different lighting conditions. Fruit images were selected in such a way that they cover a wide variety of fruit categories having same color. This could help to train a strong classifier that can detect fruit images with a wide variety.
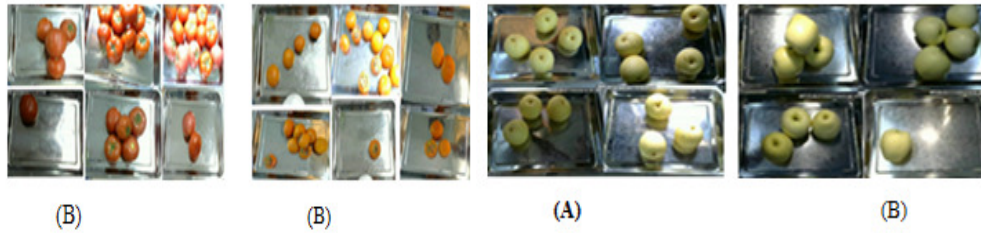
Figure2 Tomatoes (A) Persimmon (B) looks similar   Figure.3 Apple (A) and Pear (B) look similar

Figure 5 demonstrate an example of Kiwi category with different illumination parameters. Our dataset contained the samples of three different kind of kiwi fruit. The illuminations were changed due to day light exposure, it has been generated through our artificial weight machine near to the windows of our lab by opening and closing the windows, it's not artificially tampered.

.

Figure.4 Variability in the number of elements'. kiwi category.

TABLE 1 SAMPLES OF 15 DIFFERENT CATEGORIES OF FRUITS

N	Fruit Name	Total Images						
1.	Apple Categories	5024						
2.	Banana	3027						
3.	Mango	4154						
4.	Orange	3012						
5.	Peach	2629						

6.	Pear	3012	
7.	Tomatoes	2171	
8.	Guava	4008	
9.	Kiwi	4173	
10	Persimmon	2072	
11	Pitaya	2501	
12	Plum	2298	
13	Pomegranate	2167	
14	Carambola	2080	
15	Muskmelon	2078	
	Total Images	44406	

4. EXPERIMENTS RESULT & ANALYSIS

In this research analysis, we used Python 3.6, OpenCV to implement our solution. We have developed this in keras with tensor flow back-end in spyder environment which was configured to

use GPU Quadro K4200 for training. We also used pyCUDA to implement some kernels. First, we imported all keras libraries and open sources such as boken. Fruit identification in this work has been accomplished through the training of the neural network by assigning proper feature matrix and define the target class based on the input class of dataset. The database contained 44406 fruit images in which 30082 images used for training,6804 for testing &7520 for validation. All the images down-sampled to a fixed resolution of 150×150×3(length=150, width=150, & channel =3) from 320×258×3(length=320, width=258 &channel=3) and then randomly divided into training, testing and validation according to above mentioned ratio. Label representation for their respective categories is also made. Images and label are fed in to the CNN as input. Neural network used for classification is shown in figure.1. The network was trained with the stochastic gradient descent(SGD) algorithm with learning rate of 0.0001, momentum 0.9. Apparently if we carefully tune the hyperparameters for SGD and Momentum we can obtain less test error compared to Adam or RMS Prop. In my case I've found SGD optimizer work well by tuned hyperparameters properly. It was trained for 20 & 60 epochs through the training set of 30082.In the end, desired kind of fruit is determined through highest probability mechanism. In our experiments we used three Convolutional & three Maxpooling2D layer with pool- size= (2,2). At initial stage, the training loss was large which resulting in large learning rate to speed up the training process but gradually the learning rate decreased with the loss, which help in avoiding overfitting the best result. We used two drop layers with drop rate 0.5 and 0.3 to avoid overfitting of model.. After several convolutional and max pooling layers, we used fully connected layer. Fig. 5 & 6 exhibit further investigation for visual proofs by visualizing high-dimensional features space. This will give us an in depth look into their internal workings and help us in understanding them better
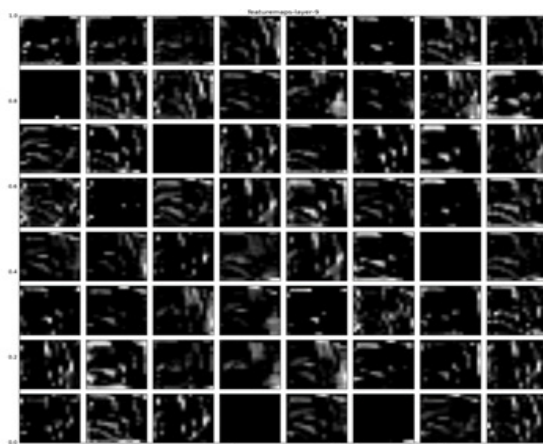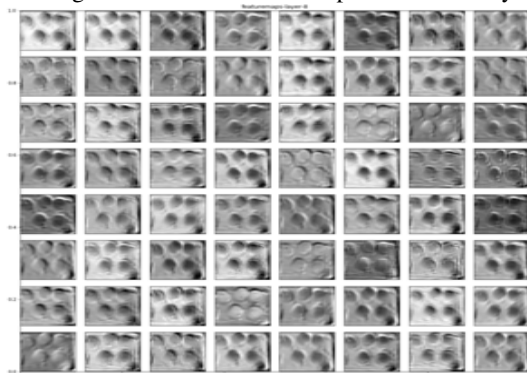

Fig.5 Features activation map from Conv8 layer

Fig.6 Feature activation map from Con9 laver

4.1 SIMULATION RESULTS

Choosing a convolutional neural network architecture for real time object detection and recognition is not an easy task because it is very difficult to estimate exact number of layers, kind of layers and number of neurons to use in each layer. In this paper, we tested different kind of CNN architecture to find the best suitable option. The CNN was first implemented without using data expansion techniques. As shown in Fig.7&8, the accuracy and loss curves are set to 20 with the training epoch during training. A large gap between training and validation accuracy occurred after 7epochs, it indicating the presence of overfitting. The highest accuracy on test images was found to be 82%, corresponding to a training accuracy 90%.
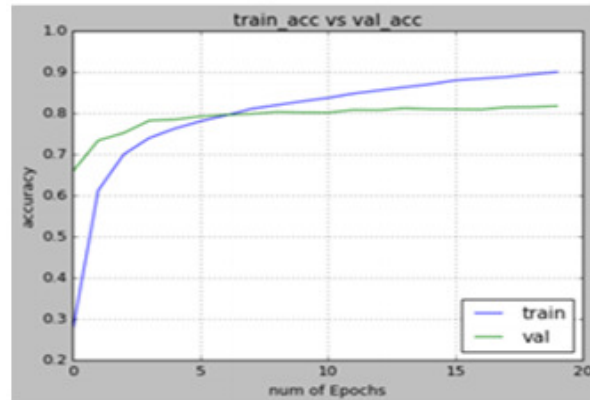

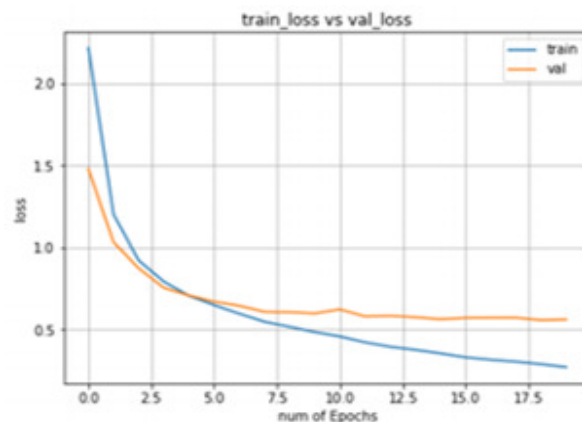Fig.7 Model validation and training accuracy curve

Fig.8 Model validation and training loss curve

From the training and validation curve in fig.7 it seemed that the CNN model accuracy could be further improved by increasing the training epochs. But when we increased the number of epochs to 60, there is no increment had been observed in validation accuracy but the training accuracy were increased to 95-96% as shown in figure 9. After repeated couple of experiments by applying different overfitting techniques like added dropout layers, used data augmentation, by testing different architectures that generalize well, varying hyper parameters, after applied these techniques there is two significant improvement were achieved. First increased the accuracy till to 99% secondly the overfitting problem removed completely. The data augmentation is one way to fight overfitting but it's not enough, so we also considered other techniques such as regularization, early stopping and all other above-mentioned techniques. Finally, we had completely removed overfitting problem as showed in fig 12.
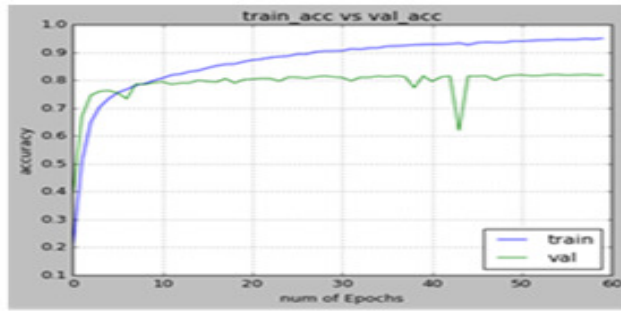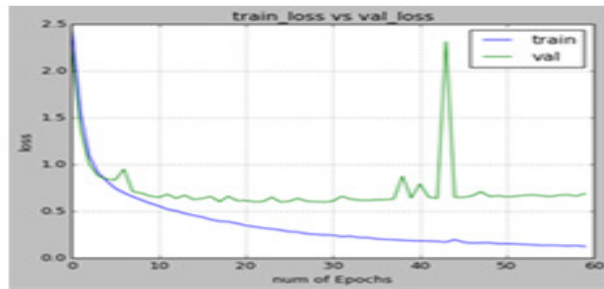
Fig.9 Validation and training accuracy curve

Fig.10 Validation & training loss curve

As shown in below figure11, the first test accuracy was greatly elevated to a level of 99% with 20 epochs and secondly the overfitting issue have been completely eliminated.
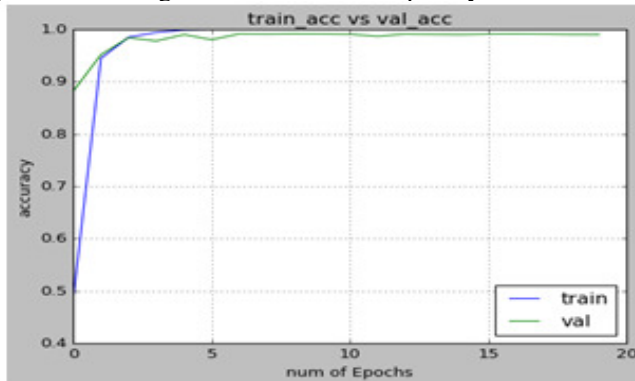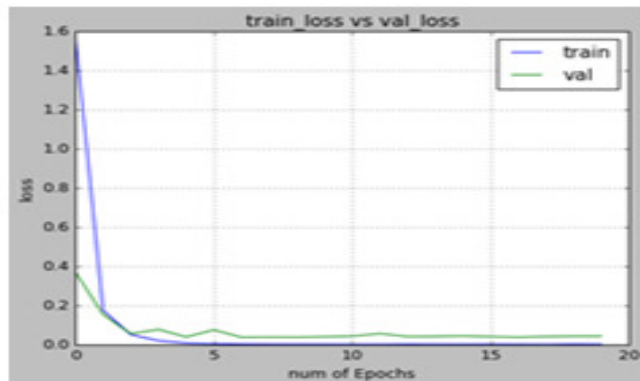

Fig.11 Model Accuracy Curves

Fig.12 Model Loss Curves

Table.2 demonstrate the classification report which presents the recognition probabilities by CNN model for test images and corresponding recognition rates for each fruit category. The precision is the ratio of tp /(tp+fp) where tp is the number of true positive and fp number of false positive. The recall is the ratio of tp /(tp+fn) where tp number of positive and fn number of false. The recall is intuitively the ability of the classifier to find all the positive samples. Support is the number of occurrence of each class in y_train. F1-score is average of precision and recall. The support represents the number of images for individual class used to test the model. We have also tested our model for new real-world dataset and overall its performance is excellent.

Table 2 Classification Report Table

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
	Avg/total	0.99	0.99	0.99	6804

4.2 CONFUSION MATRIX

Confusion matrix allowed us to look at the particular misclassified examples yourself and perform any further calculation as desired. In our confusion matrix X-axis is the predicted fruit labels and Y-axis is the true labels. The diagonal represents the correct results. We can see most of the fruit estimated correctly.
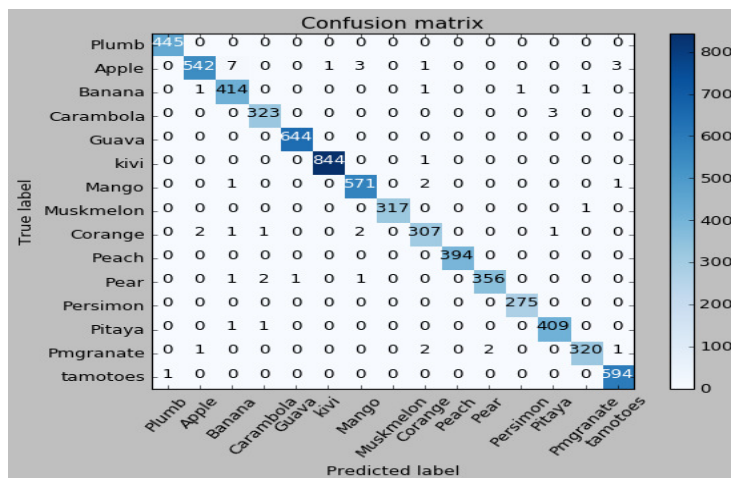

Figure 13 Confusion Matrix …………………………………..

5. COMPARISON WITH OTHER TECHNIQUES

We compared our results with several recently proposed methods [13,15,17], based on deep neural network. A fruit recognition algorithm based on convolution neural network(CNN) combined with selective search algorithm is proposed by lei Hou 2016 [13]. The recognition rate has been improved greatly up to 99.77%. But the database only contained 5330 images for 7 categories .4000 images used for training,1330 for testing. Although the recognition rate in this method is quite good but less species has been used for fruit database in that experiment. Furthermore, dataset is far from the real-world challenges and also not considering the external environment change and other factors i-e light. This is the main reason it's not more suitable for real world applications.

Control system Development for fruit classification based on Convolutional neural network has been suggested by Zaw Min Khaing in [15]. The model was evaluated under limited dataset FIDS30 as fruit image dataset consist of 971 images for 30 different kind of fruits. The proposed model had accuracy in the classification close to 94%. The model was developed using a limited dataset. So, it can't meet the actual application situation.

Fruit and vegetable classification system using image saliency and convolution neural network has been proposed in [17]. In this method the model was evaluated under limited dataset consist of 12173 images for 26 categories and have got an accuracy of 95.6%. Out of 12173 images, 80% were used for training and 20% for validation. The number of training samples are small which is not conducive to the extensibility of an algorithm.

Our proposed approach achieved almost similar accuracy with the above mention classification system but we evaluated our model in much more complicated dataset which mostly meet all real-world challenges. In our dataset many of fruits also contained their sub categories such as six different kind of apple fruits, three different type of mango and kiwi fruits, three different kind bananas fruits. Sub categories create miss matches between train and test set. Above mentioned recognition systems validate on a database which did not contain sub categories of fruit. So, it can't meet the actual application situation. We also evaluated the performance of the model under 44406 image datasets which comprised of 15 different fruit categories. Deep learning algorithms need a large amount of training samples to understand it perfectly. Hope our proposed model will adapt the real-world application requirements. In our case we have enough dataset to trained a CNN model.

6. CONCLUSION

In this paper, we proposed a novel fruit recognition classification system based on Deep Convolution Neural Network(DCNN). On the other hand, we established our own experiments database. The proposed method easily recognized fruits images with many challenges. In order to improve the functionality and flexibility, we created all real-world challenges in our dataset. Hence, the proposed method efficiently increased the recognition rate and it can meet the real word application requirements. Experiments were conducted on our own dataset and the results showed that our classification system has achieved an excellent accuracy rate of 99%. As a future direction, we can also increase the number of classes of fruits and focus on Sub-class categorization, fruit detection and localization in future work. We hope our method and database will convey more than this paper and inspire the future studies. From future perspective, more class and sub-class categorization should be included. This involve further experimenting with the structure of the network. We can also apply our algorithm in other similar fields like shrimp color monitoring.

ACKNOWLEDGMENT

CONFLICT OF INTEREST

The authors declare no conflict of interest

ABBREVIATION

The following abbreviations were used in this paper.

DCNN	Deep Convolutional Neural Network
SGD	Stochastic Gradient Decent
ILSVRC	Large-Scale Visual Recognition Challenge
ReLU	Rectified Linear Unit
RGB	Red Gray Blue
SVM	Support Vector Machine
BOF	Bag of Features
FC Layer	Fully Connected Layer
GPU	Graphic Processing Unit

REFERENCE

[1] Rocha, A.; Hauagge, D.C.; Wainer, J.; Goldenstein, S. Automatic fruit and vegetable classification from images. Comput. Electron. Agric. 2010, 70, 96–104.

[2] Bolle, R.M.; Connell, J.H.; Haas, N.; Mohan, R.; Taubin, G. VeggieVision: A Produce Recognition System. In Proceedings 3rd IEEE Workshop on Applications of Computer Vision, WACV'96, Sarasota, FL, USA, 2–4 December 1996, pp. 244–251.

[3] Hong, S.G.; Maccaroni, M.; Figuli, P.J.; Pryor, B.M.; Belisario, A. Polyphasic Classification of alternaria isolated from hazelnut and walnut fruit in Europe. Mycol. Res. 2006, 110, 1290–1300.

[4] Baltazar, A.; Aranda, J.I.; González-Aguilar, G. Bayesian classification of ripening stages of tomato fruit using acoustic impact and colorimeter sensor data. Comput. Electron. Agric. 2008, 60, 113–121.

[5] Pennington, J.A.T.; Fisher, R.A. Classification of fruits and vegetables. J. Food Compos. Anal. 2009, 22, S23–S31.

[6] Pholpho, T.; Pathaveerat, S.; Sirisomboon, P. Classification of longan fruit bruising using visible spectroscopy. J. Food Eng. 2011, 104, 169–172.

[7] Krizhevsky, I. Sutskever & G.E. Hinton, Imagenet classification with deep convolutional neural networks, in Advances in Neural Information Processing Systems, 2012.

[8] Y.Lecun, L. Bottou, Y. Bengio & P. Haffner, Gradient-based learning applied to document recognition, in Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[9] Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.

[10] K.Yanai & Y. Kawano, Food image recognition using deep convolutional network with pre-training and finetuning, in Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on, 2015.

[11] M.Oquab, L. Bottou, I. Laptev, and J. Sivic, Learning and transferring mid-level image representations using convolutional neural networks, in Proc. of IEEE Computer Vision and Pattern Recognition, 2015

[12] A Fruit Detection System Using Deep Neural Networks (Inkyu Sa, Zongyuan Ge, Feras Dayoub, Ben Upcroft, Tristan Perez, Chris McCool) Sensors (Basel) 2016 Aug; 16(8): 1222. Published online 2016 Aug 3. doi: 10.3390/s16081222

[13] L. Hou, Q. Wu, Q. Sun, H. Yang and P. Li, "Fruit recognition based on convolution neural network," 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Changsha, 2016, pp. 18-22

[14] Kaiming He et al. "Deep Residual Learning for Image Recognition." computer vision and pattern recognition (2015), pp.770-778, 2015.

[15] Z.M. Khaing, Y. Naung and P. H. Htut, "Development of control system for fruit classification based on convolutional neural network," 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), Moscow and St. Petersburg, Russia, 2018, pp. 1805-1807.

[16] Extracting Deep Features from a Single Image for Edit Propagation b Yuki Endo Satoshi Iizuka Yoshihiro Kanamori Jun Mitaniy

[17] G. Zeng, "Fruit and vegetables classification system using image saliency and convolutional neural network," 2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, 2017, pp. 613-6