VOL. 16, NO. 2, 2017, 1-4
www.fke.utm.my/elektrika
ISSN 0128-4428

ELEKTRIKA
Journal of Electrical Engineering

Convolutional Neural Network for Optimal Pineapple Harvesting

Ahmad Aizuddin Azman and **Fatimah Sham Ismail***

Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia.

*Corresponding author: fatimahs@fke.utm.my, Tel: 607-5535263, Fax: 607-5566262

Abstract: Upon ripening, colour of pineapple's peel gradually changes from green to yellowish, which spreading from bottom to the top. The objective of this project is to develop a computational intelligence method for pineapple maturity indices classification for optimal harvasting. Pineapple maturity indices can be grouped into three levels, which are unripe, partially ripe and fully ripe for determining optimal pineapple harvesting. Previous works on classifying fruit's ripeness rely on manual hand-engineered feature extraction and selection. This project proposes new intelligent method using convolutional neural network (CNN) that has the ability to learn several unique features from the given task automatically through supervised learning. The simulation results show that the method achieved 100% classification's accuracy for determining unripe and fully ripe level and 82% accuracy for partially ripe level.

Keywords: convolutional neural network (CNN); pineapple ripeness classification; python programming language;

1. INTRODUCTION

First Convolutional Neural Network has been introduced to public in early 1990 by LeCun *et al.* [1] for handwritten zip code recognition trained using backpropagation. To this date, CNN has been widely implemented for multiple task recognition task, such as action recognition [2], object recognition [3], detection of pedestrian [4], classification of traffic sign [5], face detection [6] and digit recognition [7] which have successfully show competitive result.

Previous researches on assessment of fruit's maturity has been widely done using various image processing techniques and artificial intelligence methods such as fuzzy logic and standard neural network. Specifically, previous researches rely on manual handcrafted feature extraction and feature selection based on various colour space. On contrary, CNN has the ability to learn the unique characteristic of given task automatically, which eliminates the difficulty of feature extraction and feature selection.

In Malaysia, farmers in pineapple plantation using the guidance given by Federal Agricultural Marketing Authority (FAMA) for sorting and grading the pineapple. As the sorting and grading process is done manually, it's prone to human error. On the other hand, manual inspection uses lot of man power and increase the labour cost. Therefore, by devising an algorithm using convolutional neural network to distinguish the pineapple maturity, it can be implemented on having automated pineapple harvester to assist farmer in pineapple plantation area and to monitor the quality of the pineapples.

Generally, the system designed comprises three main parts, which are input, process and output. The input is the pineapple's image, the process involves the classification by CNN and the output is the ripeness level. Before developing the CNN, the sample's preparation of different pineapple's maturity is presented. The samples are then used for training and performance's validation of the proposed CNN's architecture. Throughout the completion of project, python is used as the programming language with Keras as the deep learning library.

Pineapple ripeness can be distinguished based on the external view pell's colour before harvesting. The maturity is divided into seven stages based on the exhibited colour of pineapple's peel. Figure 1 shows the different type of pineapple maturity based on FAMA's guideline [8] for optimal harvesting. Table 1 shows the ripeness level based on the maturity indices.

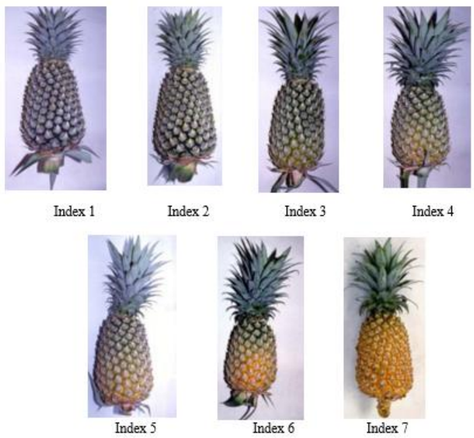

Figure 1. Different maturity indices of pineapple fruits.

Table 1. Ripeness level with corresponding index of maturity.

Level of ripeness	Index of maturity
Unripe	1, 2
Partially Ripe	3, 4, 5
Fully ripe	6, 7

2. SAMPLES PREPARATION

2.1 Collecting Samples and Labelling

Figure 2 describes the process of ripeness and index classifications. Initially, the experiment begins with preparing the sample of different maturity based on the FAMA's guideline. The samples are then cropped, only considering the region of pineapple skin without the crown and to remove the background. The samples are then resized to 200*200 pixels in size and allocated to respective ripeness level.

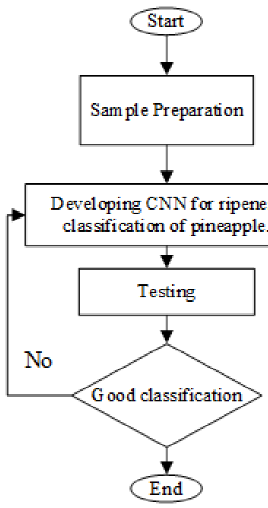

Figure 2. Experimental Process Flow.

The samples are then given associated target label based on the respective ripeness level as in Table 1. Each sample of unripe pineapple is given label [1 0 0], partially ripe pineapple is given label [0 1 0] and the last category, fully ripe pineapple is labelled with [0 0 1]. Labelling is necessary for training purpose. Basically, the CNN trying to learn unique features from training dataset of the corresponding label.

2.2 Colour Image Representation and Normalization

This project considers only RGB colour space for image representation. The raw pixel values of the image will be the input data into the network. The number of image channels are 3 for colour images. Scaling the input pixels is necessary to ease the computation and for faster convergence of the network. In this case, each pixel value in the channel is scaled down in range of [0,1] using Equation (1), given x denotes as the input pixel:

$$x = {}^{x}/_{255} \qquad (1)$$

3. CNN DEVELOPMENT PROCESS

3.1 Architecture Details

The CNN architecture includes both feature extraction and classification into a single framework. By using CNN, the design for classifier system does not need to rely on difficult hand-crafted feature extraction as the local receptive detector learns by itself through supervised learning.

The proposed architecture is inspired by LeCun [7] as shown Figure 3, comprises of input layer, two convolutional layers, C1 and C3, two non-overlapping pooling layers, P2 and P4, one fully connected layer and one output layer. The input image's size is $(200 * 200)$. C1 is composed of 8 feature maps of size $(196 * 196)$ that is obtained through convolution operation using filter of size $(5 * 5)$ with input image. Each non-overlap $(2 * 2)$ local receptive field in each feature map in C1 layer is applied with max-pooling operation, resulting in total of 8 feature map of size $(98 * 98)$ of P2 layer.

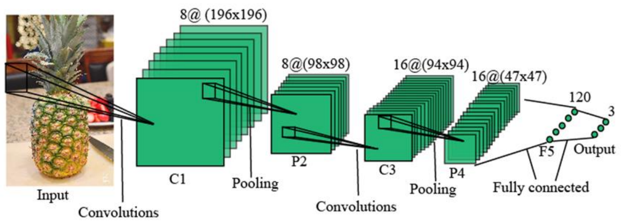

Figure 3. Le-Net4 Architecture.

C3 layer is composed of 16 feature maps of size $(94 * 94)$ resulting in convolution operation using $(5 * 5)$ filter with previous feature maps. Layer P4 has 16 feature maps of size $(47 * 47)$, which each unit in feature map of layer P4 is connected to non-overlap local receptive field of size $(2 * 2)$ in the corresponding feature map of C3.

All the multi stage feature maps are then fully connected to 120 neurons in F5 layer. Here, a dropout function with a rate value of 0.25 is introduced in F5 layer. Both convolutional layers and the fully connected layer will be passed to ReLu activation function, in order to introduce non-linearity in the network.

Neurons of F5 layer are then connected to output layer. Output layer will consist of 3 neurons corresponding to the number of classification. Softmax activation function is applied to each unit in output layer by squashing the highest output value amongst the units in output layer to '1' and suppressing the rest to '0'. Highest probability value gives the predicted class.

Basically, the CNN tries to find the correct weight value that gives the minimal error between the targeted/labelled output and predicted output for every iteration in an epoch.

3.2 Learning Process

The dataset is divided into two, which are training sample and validation sample. 90% of samples corresponding to 243 samples are used for training the network and 27 samples as validation samples. The batch size chosen is 27 for each iteration, specifically, an iteration process consist of forward propagation and backward propagation.

Throughout the training process, the weight and bias are updated frequently during back-propagation by calculating the gradient of loss function with respect to the weights in all layers (gradient descent) using normalizer algorithm. A loss function measures the discrepancy between desired output of the image and probability output of the system.

The trainable parameters suah as weights and biases are trained using back-propagation algorithm which is 'adadelta' function with learning rate 1.0 on training samples in randomized order. Hence, the gradient will update the trainable parameter 9 times for every epoch with 'categorical-cross entropy' used as the loss function.

The network will be trained for 100 epochs. For every epoch, the classifier system is tested on validation sample to observe the accuracy and loss. Early stopping during training is applied whenever the network stops learning any new features.

3.3 System Performance on Classification Design

Initially, validation samples are used to obtain the confusion matrix of the chosen network in order to visualize precision of individual ripeness level. Next, the network is tested on the random images which are not from either training dataset and validation set to determine the reliability of the network. Previously, the training and validation dataset are focusing on the pineapple's peel only, however the selected random sample test will contain the background.

4. RESULT AND DISCUSSION

4.1 Results on Learning Process

The early stopping criterion is based on the validation loss. The parameters that need to monitor are training loses and validation loses. Training loses is based on the classification performance of training dataset., and validation loses is based on the validation dataset. Figure 4 shows the graph of 'training loss versus validation loss'. The results show that if the number of epochs increase, more information is learned. This claim can be depicted by training loses as well as validation losses decreases over the epoch. However, after 30th epoch, both training loss and validation loss has slight converging. The reason to this problem is the network has stopped learning any new features.

At such state the training and validation performance should both become stationary distributions and the optimal value should occur with uniform probability anywhere between the epochs in which local optimum is reached and infinity. Hence, the training is stopped at 40th epoch to avoid overfitting.

Overfitting is the condition in which the proposed network tends to memorize the label instead of making probability. Hence, the proposed network will perform badly on sample test images which are not from training and validation dataset.

The proposed network has yielded an accuracy of 92.6% in classification of pineapple ripeness and 0.09 categorical cross entropy error. Specifically, lower categorical cross entropy error tells the difference between targeted label and predicted output is small. At this state, the proposed network has converged at certain weights and biases value.

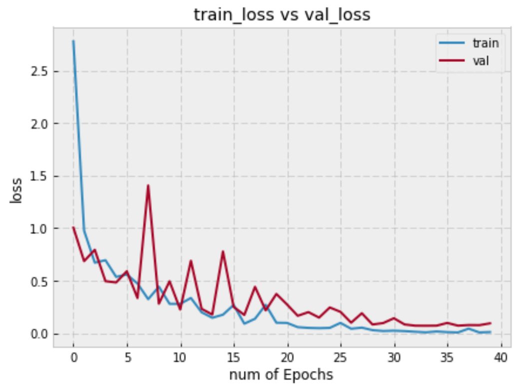

Figure 4. Training loss versus validation for 40 epochs of training.

4.1 Results on Performance Classification

The yielded accuracy of the network is based on the classification of the validation sample. A total of 27 samples consists of 8 samples of unripe category, 9 samples of partially ripe category and 10 samples of fully ripe category are used on the performance of system's classification evaluation of the proposed network.

The summary of prediction results on validation sample for RGB colour spaces are shown in confusion matrix as in Table 2. Confusion matrix is used to visualize the precision of each individual ripeness category. Here, precision intuitively describes the ability of the classifier not to label negative sample as positive.

Table 2. Confusion Matrix

Ripeness Level	Unripe	Partially ripe	Fully ripe	Precision
Unripe	8	0	0	1.00
Partially ripe	0	9	0	0.82
Fully ripe	0	2	8	1.00
Average Precision				0.94

The proposed network has lowest precision, 0.82 for determining the partially ripe category, compared to unripe category and fully ripe category. Out of 10 fully ripe samples, 2 of them is classified wrongly under partially ripe category. All the samples from unripe category and partially ripe category are categorized correctly, yielding 1.00 accuracy.

From the confusion matrix, we may say that the proposed system has outperformed [9] in classification's accuracy for unripe level with 7.5% improvement. On contrary, classification for partially ripe level has shown 1.5% increase in error. Both systems have 100% classification's accuracy of fully ripe level. Using CNN approach still produces competitive result even without the use of manual hand-engineered feature extraction and selection.

The CNN predicts the ripeness of pineapple by yielding the percentage of probability as shown in Table 3. The table shows both unripe category and fully ripe category, each category yielded 99% and 98% accuracy in average. On other hand, partially ripe category yielded 93% accuracy in average. This is due to the proposed network has lowest precision when determining the partially ripe category.

Although, the samples of test images have various background, the proposed system still able to predict the ripeness of pineapple with more than 90% accuracy. Possible reason is due to the proposed network is trained using dataset that focuses only at the pineapple's peel. So, the background's factor won't affect the prediction.

Table 3. Output probability of random sample.

Category	Sample Image	Output Probability
Unripe		Predicted Output: Unripe:[99.95396423]% Partial:[0.04603098]% Fully:[8.23526807e-06]%
		Predicted Output: Unripe:[99.46776581]% Partial:[0.5322184]% Fully:[1.44846299e-05]%
Partially ripe		Predicted Output: Unripe:[3.99420786]% Partial:[87.64442444]% Fully:[8.36136341]%
		Predicted Output: Unripe:[0.46672243]% Partial:[99.53290558]% Fully:[0.00037365]%
Fully ripe		Predicted Output: Unripe:[2.47598146e-05]% Partial:[2.97924185]% Fully:[97.02073669]%
		Predicted Output: Unripe:[1.18656814e-07]% Partial:[0.95786518]% Fully:[99.04213715]%

5. CONCLUSION

FAMA's guideline for distinguishing different maturity of pineapple has been used as the benchmarks for preparing several samples of images. Besides that, the use of Convolutional Neural Network for classifying the pineapple into 3 main categories, which are unripe, partially ripe, fully ripe has been demonstrated in this project for optimal harvesting. The analysis on the classification's performance of the proposed system showed very competitive results even without manual hand-engineered feature extraction and selection.

For future work, the designed CNN should be able to classify pineapple ripeness level in online mode. Current system only works in offline mode that's able to predict three classification levels of pineapple which are unripe, partially ripe and fully ripe by manually loading the test sample images. The online system can be implemented in an automated pineapple sorting and grading system.

The sample images for training must be at good resolution and at varying environment's condition, for example-light intensity such as morning, evening, afternoon, night at different position. Moreover, comparative performance using different colour space has to be addressed too. Comparative performance of colour space is really important if the system is intent to work in online mode. Lastly, choosing appropriate optimizer and learning rate are also necessary for smoothing the learning process.

ACKNOWLEDGMENT

The authors would like to thank for the support given to this project by Ministry of Higher Education (MOHE) and Universiti Teknologi Malaysia (UTM), under GUP grant: Q.J130000.2523.13H57.

REFERENCES

[1] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[2] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 221–231, 2013.

[3] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1106–1114, 2012,

[4] P. Sermanet and K. Kavukcuoglu, "Pedestrian detection with unsupervised multi-stage feature learning," *International Conference on Computer Vision and Pattern Recognition*, in press, 2013.

[5] D. Ciresan, U. Meier, J. Masci, and J. Schmidhuber, "A committee of neural networks for traffic sign classification," *International Joint Conference on Neural Networks*, 2011, pp. 1918–1921.

[6] M. Osadchy, Y. Cun, and M. Miller, "Synergistic face detection and pose estimation with energy-based models," *The Journal of Machine Learning Research*, vol. 8, pp. 1197–1215, 2007.

[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient based learning applied to document recognition," *Proceedings of the IEEE*, 1998, vol. 86, no. 11, pp. 2278–2324.

[8] FAMA, "Menuju kearah Kualiti Malaysia's Best Siri Panduan Kualiti Nenas", 2006.

[9] B. H Abu Bakar, A. J Ishak, R. Shamsuddin, W. Z. Wan Hassan,"Ripeness Level Classification for Pineapple Using RGB and HSI Colour Maps". *Journal of Theoretical and Apllied Information Technology*, 2013, 57, pp. 1817-3195.