Vision-Based Edge Detection System for Fruit Recognition

Sean Huey Tan1, Chee Kiang Lam1*, Kamarulzaman Kamarudin1, Abdul Halim Ismail1, Norasmadi Abdul Rahim1, Muhamad Safwan Muhamad Azmi2, Wan Mohd Nooriman Wan Yahya1, Goh Kheng Sneah3, Moey Lip Seng3, Teoh Phaik Hai3, Ong Thean Lye3

1Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis, Kampus Alam Pauh Putra, 02600 Arau, Perlis, Malaysia.
2Faculty of Mechanical Engineering Technology, Universiti Malaysia Perlis, Kampus Alam Pauh Putra, 02600 Arau, Perlis, Malaysia.
3Walta Engineering Sdn. Bhd., 13600 Prai, Pulau Pinang, Malaysia.

*E-mail: lckiang@unimap.edu.my

Abstract. There are variety of fruits around the world, different types of fruits contain different types of nutrients and vitamins which could benefits our health. In order to understand which fruit can provide specific type of nutrients, we need to identify the types of fruits. However, fruits grow in a different shape, colour and texture based on the country they were planted and the environment of the land. Implementing a machine vision-based recognition on the fruits can help people recognize them easily. In this paper, an edge detection method is applied using computer vision approach to recognize different types of fruits. The fruits are classified based on the features extracted from their images. In the experiment, a total of 450 images of three types of fruit are used, which are apples, lemons and mangoes. Pre-processing steps are applied on the captured image to improve the quality of fruit details and the edge features are extracted using Canny Edge Detection method. Classification of the fruits is accomplished using two different types of learning model, the deep leaning model, Convolution Neural Network (CNN) and machine learning model, Support Vector Machines (SVM). The performance of both classifiers is compared and the model with the best performance, SVM is chosen as the model for the system. The system can achieve 86% classification accuracy with the SVM model, which is good enough for fruit recognition.

1. Introduction
There are many kinds of fruits around the world that contains different types of nutrients [1]. In order to understand the nutrients in every fruit, the types of fruit and their origin are required to be identified. Fruit classification is one of the most common job in a fruit farm that requires plenty of workers to classify the fruits manually. Human error is relatively high in the classification process and the process also cost plenty of money [3],[4]. Therefore, in order to cut cost and increase the efficiency of fruit classification, engineers have been studying on creating a fruit recognition system [4]. The fruit recognition system makes the classification and sorting of fruit easier, and it can also be modified for automated labelling and computing price for grocery stores [4].

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Fruit recognition is a type of object recognition that requires identification of the colour and texture of the fruit [2][4]. Colour and texture are the fundamental characters of natural images that play an important role in visual perception [2]. The process of colour classification involves extraction of useful information concerning the properties on object surfaces and discovering the best match from a set of class models to implement the recognition task [5], while texture reading identifies patterns of image by extracting the dependency of intensity between pixels and their neighbouring pixels [6].

Other than the colour and texture features, the edges of a fruit are also an important feature that helps in describing the shape of the fruit. Canny edge detection method is one of the most popular image segmentation method. A Modified Canny Edge Detection method has been proposed in [3] that is believed to perform a better segmentation than the traditional canny edge detection method. The Modified Canny Edge Detection smoothen the image by computing the kernel for horizontal and vertical gradient of each pixel in the image resulting in angular direction calculated. Strong edges in images have been related to weak edges after grading.

The process of classification is the final decision-making mechanism of the recognition system. All the extracted features of the object are analysed, and the type of object will be identified by the classifier. In this project, classification of the fruits is accomplished by using two different types of learning model, the deep leaning model, Convolution Neural Network (CNN) and machine learning model, Support Vector Machines (SVM). A CNN classifier can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other [7]. CNN model with sufficient training can differentiate the aspects one from the other. Whereas SVM is a linear regression machine learning model, where its algorithm is to find a hyperplane in an N- dimensional space (N - the number of features) that distinctly classifies the data points [8].

2. Methodology
This section covers the parameters used in every step of algorithm. Figure 1 shows the flow chart of the edge detection system for fruit recognition.

![Flowchart](image)
2.1. Image Acquisition
A box with LxWxH of 38.5cmx23.5cmx22.5cm is used to setup as a platform for the image capturing of fruits. The camera uses for image acquisition is RedMi Note 4X phone rear camera with single lens CMOS sensor of 13 Megapixels resolution and 2.0 aperture. A white LED light with brightness of 100 lumens is used for illumination and a white colour cardboard is placed under the fruits as the background of the image captured. Three different types of fruit, apple, lemon and banana were captured by the phone’s camera. The fruits were placed at the centre of the base and the camera is right above of the fruits with 22.5cm measured from the base of the box. Figure 2 shows the experimental setup of image acquisition.

![Figure 2. Experimental Setup](image)

2.2. Image Cropping, Colour Space Conversion and Image Resizing
The original image of fruits captured by the Redmi Note 4X is 3120x4160 pixels with size ratio of 3:4 which is not suitable for the image processing. Hence, the captured images will be cropped into the desired size ratio which is 1:1 with 3120x3120 pixels. The cropped images which are RGB images are then converted into grayscale images by using luminosity method. The luminosity method converts the RGB images into grayscale images by forming a weighted average on red, green, and blue layers. The three weighted averages formed on red, green and blue layers are 0.299, 0.587 and 0.114 respectively. The grayscale images are then resized from 3120x3120 pixels into 512x512 pixels by performing interpolation and resampling the total pixels of the images while keeping the size ratio and the features of the images. As the size of the fruit images are reduced, lesser pixels needed to be process and hence speed up the image processing.

2.3. Canny Edge Detection
An edge detection method is used in image segmentation. Compared to the region-based segmentation method, edge detection method can extract the border better by eliminating any unwanted texture features inside the border of the fruits. Canny Edge Detection method is used due to its great noise reduction by Gaussian filter and non-maximum suppression. The Canny Edge Detector uses multi-stage algorithm to extract the edges of fruits, the process has four main algorithms which are Gaussian smoothing, computing the gradient magnitude and orientation, non-maxima suppression, and hysteresis thresholding. Figure 3 shows the block diagram of Canny Edge Detection.

![Figure 3. Canny Edge Detection](image)
A Gaussian filter is used to smoothen and reduce the noises in the fruits image as there is a lot of details in the image such as texture of fruits which are not needed in edge detection. The formula of a Gaussian distribution in 2-D is given by Equation 1:

\[ G(x|y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]  

where \( \sigma \) is the standard deviation of the distribution, while \( x \) and \( y \) represent the respective distances to the horizontal and vertical centre of the kernel. The sigma value used is 1, and the kernel sized used is 10x10.

In Sobel filter, two different kernels have been used in order to compute the gradient orientation and magnitude of the pixels in x-direction and y-direction. The size and parameters of the two kernels used are given by Equation 2:

\[ G_x = \begin{pmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{pmatrix} \quad \text{and} \quad G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{pmatrix} \]  

where \( G_x \) is the kernel for x-direction and \( G_y \) is the kernel for y-direction. Both kernels will be applied on the same fruit image neighbourhood and by applying the basic trigonometry, the magnitude of the gradient, \( G \) and the orientation \( \theta \) of the fruit image can be computed. Equation 3 and 4 show the formula for \( G \) and \( \theta \) respectively.

\[ G = \sqrt{|G_x|^2 + |G_y|^2} \]  

\[ \theta = \tan^{-1} \frac{G_y}{G_x} \]  

In order to further reduce the noises and improve the visibility of the fruit outline, non-maxima suppression is applied. Gradient magnitude of each pixel of the fruit calculated from the Sobel filter will be compared to its neighbour pixels surrounding it based on the orientations of the pixel, and by each comparison, the pixel with the highest gradient magnitude will be determined as the edge. The comparison of the pixels is separated into four groups with eight different angles which are \((0^\circ, 180^\circ), (45^\circ, -135^\circ), (90^\circ, -90^\circ), (135^\circ, -45^\circ)\).

The hysteresis thresholding is the final step of the canny edge detection, this step is to eliminate the regions which are not technically edges of the fruits but still responded as edges. The process of eliminating these regions starts by defining two thresholds, \( T_{upper} \) and \( T_{lower} \). To define the \( T_{upper} \) and \( T_{lower} \), two threshold ratios are set which are the high threshold ratio at 0.35 and low threshold ratio at 0.30. By multiplying the high threshold ratio with the maximum gradient value in each image, the value for \( T_{upper} \) for the image is obtained. While \( T_{lower} \) of the image is obtained by multiplying the low threshold ratio with the \( T_{upper} \) obtained.

2.4. Classification

Two classification models have been used to classify the pre-processed images of fruits, which are the Convolution Neural Network (CNN) classifier and Support Vector Machine (SVM) classifier. The data images for three classes, apple, lemon, and mango for both classifiers are split into two sets which are training sets and testing sets. The ratio of the data separation is 7:3 where 70% of images are used for training and 30% of images are used for validation. The total data images are 450, hence 315 images are used as training set and 135 images are used in validation set. The performance of both classifiers is compared to determine the best model.

The training images for CNN model are re-scaled at 1/255 for the model to process, the input training images undergoes random shearing, zooming and horizontal flipping to create more possibility of the training images. In CNN classifier, four 2D convolution layers with “relu” activation function have been applied into the model, the input image into the layers is scaled at 150x150 pixels dimensions. The first and second convolutions layers consist of 32 types of filtering while the third and fourth layers consist of 64 types of filtering, the filtering on four of the convolution layers are having the same mask sizes of
3x3 pixels. The 2D max pooling layers are applied after each convolution layers, and the window size used for max pooling is 2x2 pixels. The neural network is built with 4 hidden layers. The number of neurons in the first layer are 128, the second hidden layer has 64 neurons, while the third layer consists of 32 neurons, these three layers have the same activation function which is “relu” activation function. For the last layer of the neural network, which is the output layer, consists of 3 neurons which represents the three final targets of fruits.

In SVM classifier, A histogram of oriented gradients (HOG) is compiled to improve the feature extraction of fruit images for SVM. The pixels of the image will be separated into cells and the size of each cell is set at 16x16 pixels; a histogram of gradient directions is then compiled for the pixels within each cell. The HOG features extracted from the images are flattened and combined with the flattened pixels of the original image, hence obtaining a total of 262144 feature vectors for each image. Since the data of the input image are linear separable, the linear kernel is used in the SVM. In order to train the SVM, the training images is fit to the predetermined targets. The scoring for the SVM model is determined by generating predictions on the validation data.

3. Results
This section covers the procedure and the result of the entire project. In image processing part, a sample image of apple, lemon and mango is acquired from the database to illustrate the processed results. The results of each stage in image processing are shown and the observations on the sample image are explained in detail. The results of the classification model after training and validation are presented in detail with diagrams and tables.

3.1. Image Pre-processing
The images of three different types of fruits which are apple, lemon, and mango are captured by using the same image acquisition setup with a total of 450 data images. Each type of fruits is captured for 150 images in different angles, where the angle and placement of the fruits are adjusted randomly and manually. Table 1 shows the sample results of original fruit images acquired by RedMi Note4X and the processed images through each image pre-processing stages.
| Stage               | Description                                                                 | Apple | Lemon | Mango |
|---------------------|-----------------------------------------------------------------------------|-------|-------|-------|
| Original Image      | The pixel dimension of the original image is 3120x4160 pixels.              | ![Apple Original](image) | ![Lemon Original](image) | ![Mango Original](image) |
| Image Cropping      | The image is cropped into the size ratio of 1:1 with 3120x3120 pixels.       | ![Apple Cropped](image) | ![Lemon Cropped](image) | ![Mango Cropped](image) |
| Grayscale Image     | The only colour shown are the shades of grey.                               | ![Apple Grayscale](image) | ![Lemon Grayscale](image) | ![Mango Grayscale](image) |
| Image Resizing      | Difference in contrast occurs due to interpolation process.                 | ![Apple Resized](image) | ![Lemon Resized](image) | ![Mango Resized](image) |
| Gaussian Filter     | The image is blurred and becomes brighter.                                 | ![Apple Gaussian](image) | ![Lemon Gaussian](image) | ![Mango Gaussian](image) |
| Sobel Filter        | Strong edges are extracted but consists of noises.                          | ![Apple Sobel](image) | ![Lemon Sobel](image) | ![Mango Sobel](image) |
| Non-maxima suppression | Noises reduced but the edges become thinner.                     | ![Apple Non-maxima](image) | ![Lemon Non-maxima](image) | ![Mango Non-maxima](image) |
| Hysteresis Thresholding | The edges are highly visible with very little noises inside the border. | ![Apple Hysteresis](image) | ![Lemon Hysteresis](image) | ![Mango Hysteresis](image) |
3.2. Classification

The performance of CNN and SVM classifiers are compared based on confusion matrix and classification report as shown in Table 2. The CNN classifier is a sequential model, while the SVM is a linear regression model. According to the same number of training images and testing images, the F1-scores of SVM classifier for all three fruit classes are much better when compared to that of CNN classifier. In the apple class, SVM scores a high value of 0.80 over 1.00, while CNN only scores 0.45. The CNN classifier scores the lowest in lemon class with only 0.24, while SVM have the highest score on lemon class compared to the other classes of its own with 0.90. In the mango class, although CNN scores the highest in this class compared to other class of its own, it is still a very low score at 0.46, while the performance of SVM is very stable as it scores 0.88 in mango class. The overall accuracy of CNN model and SVM model is 40% and 86% respectively. It is clearly shown that the performance of fruit classification by SVM classifier is dominant over the CNN classifier in all aspects.

Table 2. Comparison of CNN classifier and SVM classifier

| Characteristics          | CNN Classifier | SVM Classifier |
|--------------------------|----------------|----------------|
| Type of model            | Sequential     | Linear         |
| Total number of training images | 315            | 315            |
| Total number of testing images | 135            | 135            |
| F1-score of apple class  | 0.45           | 0.80           |
| F1-score of lemon class  | 0.24           | 0.90           |
| F1-score of mango class  | 0.46           | 0.88           |
| Model Accuracy           | 40%            | 86%            |

4. Discussion

Consequently, with the total number of 450 data images used, the SVM model is more suitable than CNN model to perform the fruit recognition. Although the CNN classifier is a deep learning model which is easier to be built and able to perform training more efficient, it requires a large amount of database to train the model and multiple algorithms to tune the model for higher accuracy. In contrast, the SVM classifier is slower in model training and testing when compared to CNN classifier, it also requires feature extraction manually before inputting the image into the model. In this case, there is only one feature that has been extracted from the fruit images, which is the edge of the fruits. Since the shapes of apple, lemon, and mango have certain similarities, for the CNN model to classify these fruits accurately, it requires a huge amount of fruit images for training to learn the difference in edges of the fruits. While for the SVM classifier, a new feature is created for the fruit images which is the HOG features. As a result, the model can differentiate the three classes of fruit better and hence achieving higher performance accuracy.
5. Conclusion
In conclusion, the research work on applying the edge detection method in fruit recognition system has been successfully conducted. The edge for three classes of fruits, apple, lemon, and mango have been extracted and used as input data images for the classification models. Both the CNN classifier and SVM classifier can produce results using the fruit’s edge images as input.

The CNN classifier and SVM classifier are applied for the fruit recognition system. The performances of both classifiers have been compared with details in characteristics. The model accuracy of SVM classifier (86%) is much higher compared to the model accuracy of CNN classifier (40%). Hence, the SVM classifier is a more suitable model for the fruit recognition system.

The fruit recognition system based on extracted edges is successfully developed. There are three types of fruit that can be recognized by the system such as apple, lemon and mango. The SVM model has been chosen as the main classifier for the fruit recognition system because it has a better performance in overall model accuracy and better performance in prediction accuracy for each class of fruits. With a model accuracy of 86%, the SVM classifier can achieve good performance in fruit recognition system.

Acknowledgements
The authors would like to acknowledge the support from the University-Private Matching Fund (UniPRIMA) from the Research Management Centre, UniMAP and Walta Engineering Sdn. Bhd.

References
[1] Pham S H 2006 Importance of fruits and vegetables in cancer prevention, Nutritional Perspectives: Journal of the Council on Nutrition 29(3) pp 37-46
[2] Arivazhagan S, Shebiah R N, Nidhyanandhan S S and Ganesan L 2010 Fruit recognition using color and texture features, J. Emerging Trends Comput. Inf. Sci. 1(2) pp 90–94
[3] Monir Rabby M K, Chowdhury B and Kim J H 2018 A Modified Canny Edge Detection Algorithm for Fruit Detection & Classification, 10th International Conference on Electrical and Computer Engineering (ICECE), (Dhaka, Bangladesh) pp 237-240
[4] Mhaski R R, Chopade P B and Dale M P 2015 Determination of ripeness and grading of tomato using image analysis on Raspberry Pi, Communication, Control and Intelligent Systems (CCIS) (Mathura) pp 214-220
[5] Sahin F 1997 A Radial Basis Function Approach to a Color Image Classification Problem in a Real Time Industrial Application (Master’s thesis, Virginia polytechnic institute, Blacksburg)
[6] Kartikeyan B and Sarkar A 1991 An identification approach for 2-D autoregressive models in describing textures, Graphical Models and Image Processing 53 pp 121-131
[7] Kumari R S S and Gomathy V 2018 Fruit Classification using Statistical Features in SVM Classifier, 4th International Conference on Electrical Systems (ICEES) pp 526-529
[8] Li Q, Cai W, Wang X, Zhou Y, Feng D D and Chen M 2014 Medical image classification with convolutional neural network, 13th International Conference on Control Automation Robotics & Vision (ICARCV) pp 844-848