Classification of Green Coffee Beans by Convolutional Neural Network and its Implementation on Raspberry Pi and Camera Module

Hidekazu Fukai [a], Junya Furukawa [a], Hiroki Katsuragawa [a], Carlito Pinto [b] and Carmelita Afonso[c]

[a] Faculty of Engineering, Gifu University, Gifu 501-1193, Japan
[b] Department of Informatic Engineering, National University of Timor Lorosa'e, Hera-Dili, Timor-Leste
[c] Alumni of Department of Informatic Engineering, National University of Timor Lorosa'e, Hera-Dili, Timor-Leste

E-Mail: fukai@gifu-u.ac.jp

Abstract: The coffee is the most important agricultural product of Timor-Leste for the acquisition of foreign currency. Nevertheless, there are almost no rationalizations at local production sites. The efficient enhancement of the value of coffee beans is desired. The grade of a lot of coffee beans greatly depends on the number of defect beans. However, the defect beans are currently removed by hand pick in Timor-Leste. The final objective of our study is to develop the automatic coffee beans sorting system for the producers of coffee beans in Timor-Leste using state-of-the-art machine learning techniques and cheap single-board computer. As the first step, we developed an image processing system which classifies the images of green coffee beans into each type of defect by deep convolutional neural networks. Next, the trained artificial neural network was implemented into Raspberry Pi compute module with camera module. We evaluated the performances of inspection speed and accuracy of the sorting system for practical realization.

Keywords: Agricultural engineering, Manufacturing automation, Convolutional neural networks, Image processing, Raspberry Pi

A. Introduction

In Timor-Leste, the coffee is an important crop that assures a sustainable economy to farmers and the principal income for 1/4 of the Timorese population. The coffee is the main export commodity of Timor-Leste [1]. According to the International Coffee Organization (ICO), Timor-Leste produced 80,000 of the 60 kg bags coffee in 2016 [2]. Although coffee producers are located throughout Timor-Leste, concentrated particularly in Ermera, Aileu, Ainaro, Manufahi and Liquica. Traditionally, Timor-Leste's coffee producers have processed their own coffee totally from cherry on the bush to commercial green beans. To compete in the world market of coffee beans, not only the amplification of production power but also improvement of the quality of coffee beans is essential. Apart from the type of subspecies of coffee beans, the grade of a lot of green coffee beans greatly depends on the number of defect beans. There are several types of defects such as black bean, broken bean, stones, etc. which affects their grade.

Generally the quality inspection of coffee beans is still performed by human experts in Timor-Leste [3]. This manual sorting by visual inspection is labor intensive, time-consuming and suffers from the problem of inconsistency and inaccuracy in judgment by a different human. Therefore, introduction of automatic coffee beans sorting systems is desired to push up the value of Timor-Leste's coffee products and to improve international competitiveness.

There are already several commercial general-purpose hardware systems which sort defect crops or beans automatically. For example, SATAKE Corporation (Japan) produced full color optical sorter, PIKASEN FMS200, in 2014, which can be applied for green coffee beans [4]. However, these commercial systems are expensive and usually difficult to introduce into local coffee bean production site of under developing countries. Moreover, usually the commercial general-purpose sorting systems sort crops into two classes, i.e. good or no good. On the other hand, in coffee beans case, there are several types of defects and the deduction points are different depending on them on quality assessment of a lot of coffee beans (Ch. II-A), so the development of sorting system which can recognize each type of defects, and control each accuracy of sorting is desired.

The final goal of this study is to develop an automatic defect coffee beans sorting system which recognize several types of defects and also can control the accuracy of sorting for each defect type. Also, to install the system in rural coffee bean association in Timor-Leste, it is better if the system could consist of cheap modules. For this objective, as the first step, we developed a computer program whose input is color image of green coffee bean and output is probability of each category of defect. As the estimation system, we used deep convolutional neural networks, a state-of-art technique of the image processing field [5].

In this study, we also applied conventional support vector machine (SVM) with linear kernel for the classification and the performances of CNN and SVM were compared in terms of the accuracy of classification, calculation speed. Next, the trained neural network was installed into Raspberry Pi compute module. The Raspberry Pi is a tiny and affordable single-board computer with operating system and several I/O ports. The Raspberry Pi is suitable for our project because we can easily obtain the module in worldwide in cheap price and we can easily control the camera module and other devices. We evaluated the feasibility of the sorting system using Raspberry Pi from the viewpoint of accuracy and data processing time for each image.

This work is conducted in collaboration of the faculty of engineering of National University of Timor Lorosa'e (UNTL) and the faculty of engineering of Gifu University of Japan, supported by Japan International Cooperation Agency (JICA for developing countries under the project for capacity development of faculty of engineering, science and technology, the national university of Timor-Lorosa'e (UNTL) Phase 2 (CADEFEST2).

B. Materials and Methods

A. *Coffee Beans*

The quality of coffee beans depends on techniques of ripe cherries collection, removal of the fruit layers by dry or wet process, moisture reduction, and so on. These technical factors affect the physical quality of coffee beans, i.e., the uniformity of beans, size and shape, the presence of foreign matter, immature beans, damaged beans [6][7]. According to the International Coffee Organization [2] and the Specialty Coffee Association of America (SCAA) [8], 300g of green coffee beans can admit 86 defects for Arabica varieties and 150 defects for Robusta at most as good quality class. The uniformity of the beans (size, shape and color) is important in the roasting and industrial processing because the best roasting condition and procedure such as temperature and time differs depending on each type of bean [6]. The color of coffee beans also has a significant influence on the price in the market [9]. The blue color to the gray-green of the beans is the most desirable color. The brown to black, light brown to dark brown, yellow semi-transparent, amber to yellow are undesirable color (or defect bean color). A non-uniform coloration of beans indicates non-uniform quality.

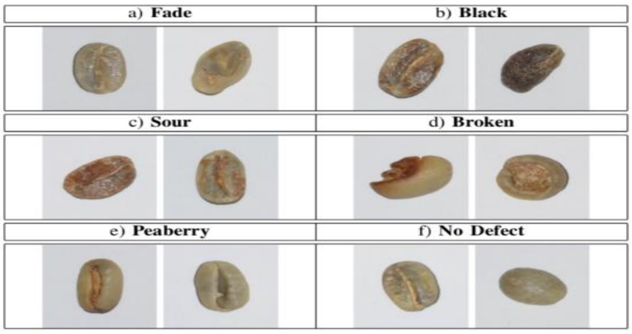
There are several defect types on green coffee beans. The number of deduction points are different according to the defect types (Table I) because the contamination of defect beans affects the taste of coffee depending on the defect types. The value of coffee beans depends on the total number of deduction points in a certain amount of beans. It is important to remove defect beans paying attention to the deduction point. The photo examples of each type of beans used in this study are shown in Table II. The specifics of each type of beans are described in the following.

TABLE I
EXAMPLE OF DEDUCTION POINTS OF THE DEFECT BEANS

Type	Deduction point
Fade	0.2
Black	1
Sour	0.5
Broken	0.2
Insect Damage	0.2 to 0.5
Stone, Piece of Wood	1 to 5

Example of deduction points of defect beans in Brazil. Same table is used in Timor. It describes deduction point for one bean.

TABLE II
EXAMPLES OF BEANS



a) Fade: Fade is unroasted coffee beans that have lost much of their original color. This is a characteristic of old crops and beans that were dried too rapidly. Processed coffee beans will slowly fade from green to pale yellow, if stored too long before roasting. It is also called ”soapy” or ”bleached”.

b) Black (Black or partial black bean): Black is black, or very dark, unroasted beans. This is typical result from harvesting immature cherries or harvesting dead cherries that fall naturally from the tree. This can also result from exposure to water or heat, or from insect-damage. Unroasted coffee beans with more than 25% black, deep blue, or dark brown surface area, may be considered black beans. Black beans have a detrimental effect on coffee taste. The number of black beans in a representative sample is a basic measure of coffee grade.

c) Sour (sour or partial sour bean): Sour can be caused by overripe or fermented cherries, or by improper fermentation. Sour beans are recognized by a yellow or yellowish-brown to reddish-brown color. It is often possible to see a dark or black spot and beans still covered by silver skin. When cut or scratched, a sour or vinegar-like smell is released. The sour beans are caused by the death of its internal embryo, which may result from over-fermentation and high temperature at multiple points during harvesting and processing, over- fermentation in the fruit still attached to trees under humid conditions and drying up at high temperatures.

d) Broken (Cut/nipped bean and pressed or crushed bean): Broken is wet processed beans that are cut or bruised during pulping. It is typically caused by damaged or improperly configured pulping equipment. Pulper cut beans will usually show brown or black marks after processing. Dis- coloration develops by oxidation at the damaged areas and off-flavors may result. Pulper damaged beans roast unevenly, age rapidly, and are susceptible to damage by vapors, dust, and other adverse environments. It is also called ”blackish” or ”pulper cut”.

e) Peaberry (Small and peaberry): A single rounded bean from a coffee cherry which bears one bean instead of the usual flat sided pair of beans. It also known as ’caracol’, ’perla’ and ’perle’. On the roast of coffee beans, beside the uniformity of size, the uniformity of shape is also preferred for the uniformity of roast so the peaberry should be separated. The peaberries are frequently separated and sold as a distinct variety with special value.

f) No Defect: This is the bean which have no defects. It is not a type of defect, but we call it no defect type for convenience.

B. Data Acquisition

The inputs of neural networks were digital images of green coffee beans. The samples of coffee beans were from Timor- Leste, harvested in 2016. The green coffee beans were placed on white paper and the images were taken with Nikon digital camera. The camera was set in automatic mode with F/16 f-number, exposure time 1/60 s, ISO 200, exposure compensation 1.3, and autofocus mode. The camera was put at 1 m above the surface of the beans. The both front and back sides of the beans were taken. Three lighting devices were placed around the beans to make the brightness around the beans uniform. The shooting environment is shown in Fig. 1 and Fig. 2. Image preprocessing techniques were applied to the photographs to isolate each bean (Fig. 3). Totally about 13,000 full color images were obtained from 6,500 beans. The size of each image was set to 64×64 pixels. The images were manually labeled as fade, black, sour, broken, peaberry and no defect. Note that one image may have multiple labels (e.g. A bean was labeled to broken and peaberry). The images were divided into three groups, training data, validation data and test data to apply the machine learning techniques. The training data was used for training the machine learning models. The validation data was used to confirm the transition of classification performance in the training phase of the neural networks. The test data was used to evaluate the sorting ability of the trained models. The number of data for each group is shown in Table III.

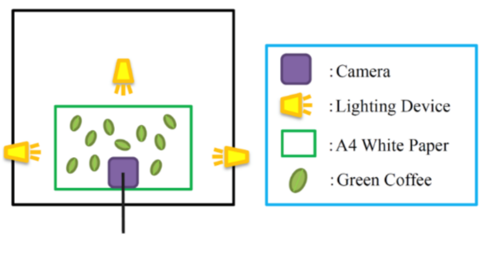


Fig. 1. Photographic environment from the above.

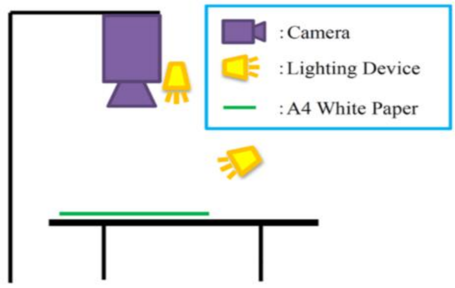


Fig. 2. Photographic environment from the side.

Fig. 3. Images obtained by image processing.

C. *Convolutional Neural Network*

The convolutional neural network (CNN) is a type of the artificial neural network which is mainly applied for image recognition [10]. The CNN consists of three kinds of layers (Fig. 4). The convolution layer has the role of extracting features of image using spatial filters. Assume that the input from the (l−1)th layer proceeds to the lth layer with convolutional connection. Let the input has the form $W\times W\times K$ and the convolution layer has M types of spatial filters in the form $H\times H\times K$. Then, the output $uijm$ is calculated by convolution of input and this spatial filter as eq. (1).

TABLE III
THE NUMBER OF DATA FOR EACH TASK

Task	Training	Validation	Test
Fade	4096	584	1170
Black	1146	162	326
Sour	4096	584	1170
Broken	308	44	88
Peaberry	1940	276	554

The task 'Fade' means the binary classification of 'Fade beans' and 'No defect beans'.

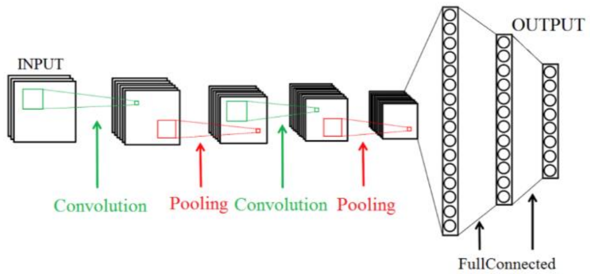


Fig. 4. Typical structure of CNN.

$$u_{ijm} = \sum_{k=0}^{K-1}\sum_{p=0}^{H-1}\sum_{q=0}^{H-1} z_{i+p,j+q,k}^{(l-1)} h_{pqkm} + b_{ijm} \qquad (1)$$

Here, the parameter $bijm$ is a bias of the layer. Then, the activation function $f(\cdot)$ is applied to the $uijm$ as eq. (2) and we will get $z_{ijm}{}^{(l)}$ as the output of the convolution layer.

$$z_{ijm}^{(l)} = f(u_{ijm}) \qquad (2)$$

The shape of the output changes depending on the stride of the filter. For example when the stride value s is 2, width and height of the output will be half of the input. A schematic diagram of the convolution is shown in Fig. 5.

Pooling layer is usually placed just after the convolution layer. The role of pooling layer is to decrease position sensitivity of the feature which extracted by the convolution layer, so even if the position of the target feature in the image shifts slightly, the output of the pooling layer does not vary. There are several types of pooling methods. Max pooling was used in this study. The max pooling is a method which sets the maximum value contained in Pij as an output value $uijk$ described as eq. (3).

$$u_{ijk} = \max_{(p,q)\in P_{ij}} z_{pqk} \qquad (3)$$

The max pooling is applied independently for each channel k. The output size depends on the stride as same as convolution layer.

Fully connected layer has a role of converging features obtained by repeating convolution and pooling to the number of classes which we want to classify. This is same architecture used in the classical artificial neural networks. The output of the full connected layer zj is given by eq. (4) and eq. (5).

$$u_j = \sum_{i=1}^{I} w_{ji} x_i + b_j \tag{4}$$

$$z_j = f(u_j) \tag{5}$$

Here $f(\cdot)$ is the activation function.

In this study, rectified linear unit (ReLU) function is used for the activation function for the convolution layers and the fully connected layers. The ReLU function is described as eq. (6).

$$f(x) = \begin{cases} x \ (x \geq 0) \\ 0 \ (x < 0) \end{cases} \tag{6}$$

In final layer, softmax function is used. The softmax function is generally used to get the output value of each unit as the probabilities that the input belongs to each class. The softmax function is described as eq. (7).

$$f(x) = \frac{e^x}{\sum_{k=1}^{K} e^x} \tag{7}$$

The training of the network was achieved using GTX1080 GPU.

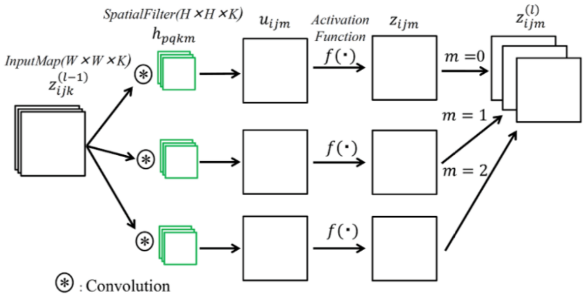


Fig. 5. A schematic diagram of the convolution.

D. Raspberry Pi and Camera Module

Raspberry Pi is a series of credit card size single-board compute modules developed by the Raspberry Pi Foundation to promote the teaching of basic computer science in schools and in developing countries [11] (Fig. 6). The CPU of Raspberry Pi is ARM processor and several types of operating system can be run on the module. The Raspberry Pi is suitable for our project because we can easily obtain the module in worldwide at low price. We can also control the general camera module and other devices easily.

In this study, Raspberry Pi 3 Model B was used. It has a 1GB RAM, a quad-core 64-bit ARM processor, a HDMI port, 4 USB ports and wired LAN port, and supports Wi-Fi and Bluetooth wireless connection. Various additional modules can be connected directly to the base, e.g. camera, display, various sensors and so on. It also has a GPU and can be manipulated by simple LCD. In addition, it has

TABLE IV
PARAMETERS OF THE CNN

Layer	**Filter size** $(H \times H \times K)$	**Stride** (s)	**Output map size** $(W \times W \times M)$	**Activation function** $f(\cdot)$
Input	-	-	$64 \times 64 \times 3$	-
Convolution1	$3 \times 3 \times 3$	1	$64 \times 64 \times 32$	ReLU
Pooling1	$2 \times 2 \times 32$	2	$32 \times 32 \times 32$	-
Convolution2	$3 \times 3 \times 32$	1	$32 \times 32 \times 64$	ReLU
Pooling2	$2 \times 2 \times 64$	2	$16 \times 16 \times 64$	-
Convolution3	$3 \times 3 \times 64$	1	$16 \times 16 \times 128$	ReLU
Pooling3	$2 \times 2 \times 128$	2	$8 \times 8 \times 128$	-
Convolution4	$3 \times 3 \times 128$	1	$8 \times 8 \times 256$	ReLU
Pooling4	$2 \times 2 \times 256$	2	$4 \times 4 \times 256$	-
FullConnected1	-	-	$1 \times 1 \times 512$	ReLU
FullConnected2	-	-	$1 \times 1 \times 2$	Softmax

The top row represents the first layer.

46 pin general purpose input / output terminal (GPIO). The GPIO makes us possible to use electronic devices and handle various sensors.

As the camera module, we used original Raspberry Pi Camera Module V2 which is connected to CSI-2 connector.

Fig. 6. Raspberry Pi 3 Model B

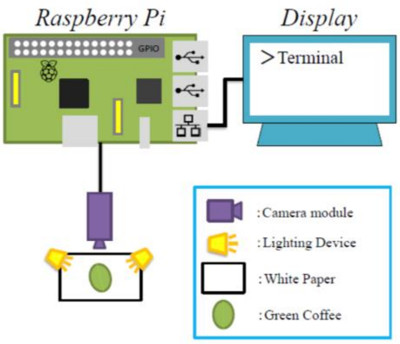


Fig. 7. Implementation of the classifier on Raspberry Pi

The Camera Module V2 has a SONY IMX219 8-megapixel sensor and can be used to take high-definition video, as well as still photography. We can use the libraries bundled with the camera. It supports 1080p30, 720p60 and VGA90 video modes, as well as still capture. The camera module was attached to the CSI port on the Raspberry Pi via 15cm ribbon cable. The reason why we used the original camera module instead of USB third party cameras is that CSI-2 connector is faster than USB 2.0.

3. Experiment

A. *Classification of Green Coffee Beans using CNN and SVM*

There are several defect types which we want to label on green coffee beans. Generally, the softmax function is used as the output layer for multiple classifications because its outputs can be considered to be values of probability of each class. However, in the case of coffee beans, some beans have multiple defects. For this reason, we performed five binary classifications for each type of defect beans with no defect beans. The number of bean images used for each classification task is shown in Table III. The parameters of CNN are shown in Table IV. layer [12] was added after 'FullConnected1' layer and set dropout rate to 0.5. One of the deep learning libraries Chainer [13] was used to implement the CNN.

We also applied conventional linear SVM instead of CNN for comparison of the performance in accuracy of classification and processing time. On implementing linear SVM, the parameter C was set to 1 and parameter gamma was set to 50. One of the machine learning library scikit-learn was used to implement the SVM. In the SVM tasks, the data was reshaped to one dimensional vector form to adapt to the SVM. We also applied nonlinear kernel SVM but the performances of accuracy were almost the same with linear ones so we will discuss about the linear ones.

B. *Implementation of CNN on Raspberry Pi*

A After the training of CNN parameter on desktop PC, we implemented the coffee bean classifier of trained CNN models on Raspberry Pi. Python (a programing language), OpenCV (an Open Source Computer Vision Library), and

Chainer (an open source library of neural networks for deep learning) were installed on Raspberry Pi. We developed a python program which takes a photo of a green coffee using the camera module, converts the photo image with OpenCV to fit the image as an input of CNN, and makes classification of the types of the coffee bean image using CNN with pretrained network parameters. The schematic illustration of the system is shown in Fig. 7. We estimated the performances of the classifier on three types of binary classification tasks of Black, Sour and Peaberry from good beans. We also checked the processing speed of the classifier. Because the power of the Raspberry Pi CPU is limited and calculation time is heavily depending on the image size, we prepared 4 trained CNN models for four kinds of image size: 32×32, 64×64, 128×128 and 256×256. We measured the mean processing time for classification of one bean in each image size and also the accuracy of classification.

4. Results

The five types of binary classifications were performed by both CNN and SVM on desktop PC. The values of classification accuracy are shown in Fig. 8. We got better performance from CNN than SVM in all types of classification. We got high accuracies (over 90%) in the classification of Black and Sour. The classification accuracy of Fade, Broken and Peaberry were in the 75% to 90%, which seems not enough performance for the practical sorting system.

We checked the performance of the classifier implemented on Raspberry Pi. The comparison of classification accuracies by the CNN and the Raspberry Pi classifier is shown in Fig. 9.

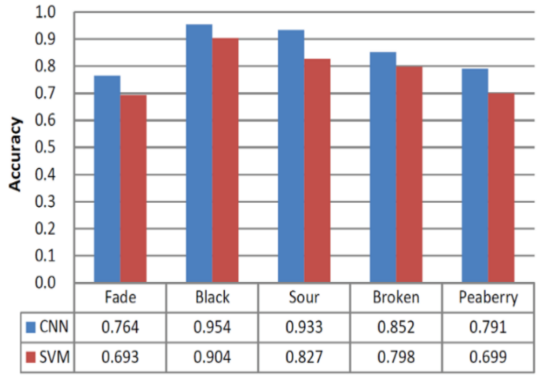


Fig. 8. Accuracies of CNN and SVM classification.

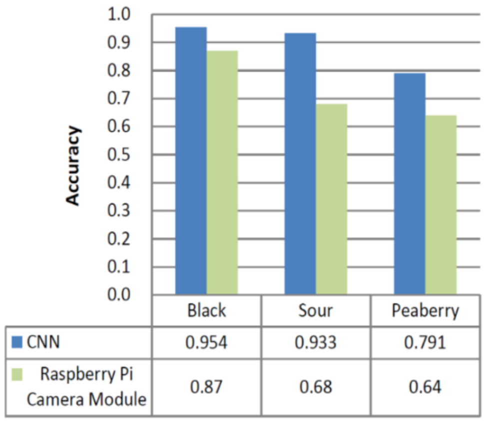
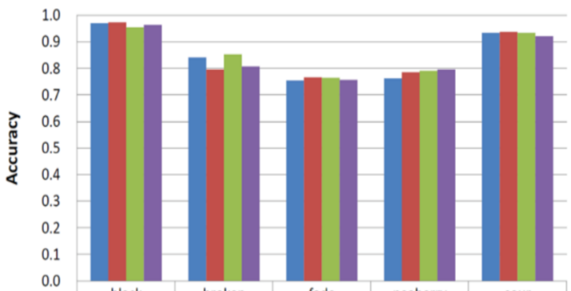


Fig. 9. Accuracies of the CNN and Raspberry Pi classifier.

We also measured the processing time of classification of one bean by the Raspberry Pi system with four kinds of image size, using network parameters which were pre-trained on desktop PC in advance. As the results shown in Fig. 10, the accuracies of classifications were not affected by the image size so much. On the other hand, the processing time were dramatically decreased with smaller size of images (Fig. 11). It was about 0.6 second on average for the classification of one bean when we used 32 × 32 pixels image.

5. Discussion and Conclusion

As the first step of the development of coffee bean sorting system, we classified five types of defect beans using Deep Convolutional Neural Networks (CNN). The results of CNN were compared with the results of conventional linear



	black	broken	fade	peaberry	sour
256×256	0.969	0.841	0.754	0.762	0.933
128×128	0.972	0.795	0.766	0.785	0.937
64×64	0.954	0.852	0.764	0.791	0.933
32×32	0.963	0.807	0.756	0.796	0.921

Fig. 10. CNN classification accuracies for several image size.

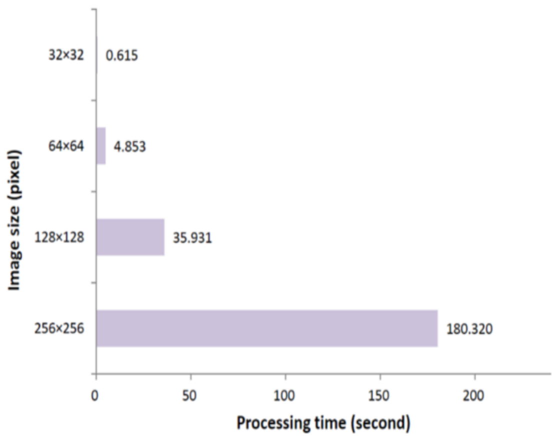


Fig. 11. The processing time to classify each image size.

SVM. We got better performance with CNN in all the defect types.The classification accuracy were more than 90% on classification for Black and Sour beans with CNN, yet the accuracies for Fade, Broken and Peaberry were in 75% to 90%.

It is well known that the deep learning requires huge number of data for training. However, generally it is difficult to prepare training data especially of defect samples. One reason why we could not get good performance on some types of defect beans in this study may be the lack of size of training data. Even if we want to prepare a lot of samples, if the defect type is rare, it is difficult to prepare enough amounts of samples for training the network. The performance could be improved by preparing more samples for training, reconsidering the structure of CNN, or artificial data augmentation by such as random rotation, shifts, and flips.

In addition, we implemented the trained classifier on Raspberry Pi and examined its performance to check its feasibility in terms of accuracy of classification and processing speed. As result, the system could proceed whole the process from taking a picture to getting classification result for one bean by 0.6 second when the processing image size was 32 × 32 pixels. Because the processing time is greatly affected by the image size, yet the performances of classification accuracy were not so affected. We need to specify the optimized parameter of image size in future study.

The reason of the difference in the performances of classification accuracy between digital camera and Raspberry Pi camera module is that we used the pre-trained network the Raspberry Pi system. On the other hand, it is not realistic to train the network parameters on Raspberry Pi itself because of poor CPU and GPU performance. To improve the accuracy of classification acquire the better trained network for Raspberry Pi, we need to specify the differences of images which are taken by digital camera and Raspberry Pi camera module. We can also consider taking enough amounts of sample images for training by Raspberry Pi camera module first and train the network with high performance CPU and GPU, then reinstall the network parameter into Raspberry Pi system.

In terms of feasibility of the development of coffee beans sorting system with Raspberry Pi and CNN, the processing time of 0.6 second or less for one bean seems enough if the machine is built with several Raspberry Pi and camera modules arranged in parallel.

References

[1] M. Ximenes, “A Teconologia Pos-Colheita e Qualidade Fisica e Organoleptica do Caf Arabica de Timor,” Dissertac ̧a ̃o para obtenc ̧a ̃o do grau de mestre em Engenharia Alimentar, ISA-UTL, Lisbon Technical University, Portugal, 2010.

[2] International Coffee Organization, “Total production by all exporting countries”. [Online] Available URL: http://www.ico.org/prices/po-production.pdf

[3] L. Gomes, “Efeito de Diferentes Tcnicas de Secagem na Qualidade do Caf Arabica (Coffea Arabica L.) em Timor-Leste,” Dissertac ̧a ̃o para obtenc ̧a ̃o do grau de mestre em Engenharia Agronomica., Universidade de Evora Escola de Cieˆncia e Tecnologia, Evora, Portugal, 2014.

[4] Satake USA, Inc., PIKASEN FMS2000, [Online] Available URL:http://www.satake-usa.com/pikasen-alpha/

[5] C. Pinto, J. Furukawa, H. Fukai, and S. Tamura, “Classification of Green Coffee Bean Images Based on Defect Types Using Convolutional Neural Network (CNN)”, Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), 2017 International Conference on. IEEE, pp.1– 5. 2017.

[6] A. Franca, J. Mendonca, and L. S. Oliveira, “Composition of green and roasted coffees of different cup qualities,” LWT vol. 38, pp.709-715, 2005.

[7] D. J. Kwon, H. J. Jeong, H. Moon, H. N. Kim, J. H. Cho, and J. E. Lee, “Assessment of green coffee bean metabolites depend on coffee quality using H MNR based metabolic approach,” Food Research International, vol. 67, pp.175-182, 2015.

[8] Specialty Coffee Association of America, [Online] Available URL:http://www.coffeeresearch.org/coffee/scaaclass.htm

[9] F. M. Borem, F. C. Ribeiro, L. P. Figueiredo, G. S. Giomo, V. A. Fortu- nato, and E. P. Isqueirdo, “Evaluation of the Sensory and color quality of the coffee beans stored in hermetic packaging”, Journal of Stored Product Research, vol. 52, pp.1–6, 2013.

[10] Y.LeCun,L.Bottou,Y.Bengio,andP.Haffner,“Gradient-basedlearning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp.2278–2324, 1998.

[11] [Online] Available URL:https://www.raspberrypi.org/

[12] N. Srivastava, G. E. Hinton, A.Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. JMLR, 2014.

[13] [Online]Available: URL:https://chainer.org/