Convolutional Neural Networks Implementation for Chili Classification

Tuti Purwaningsih, Imania Ayu Anjani, Pertiwi Bekti Utami
Department of Statistics
Islamic University of Indonesia
Yogyakarta, Indonesia
tuti.purwaningsih@uii.ac.id, 14611190@students.uii.ac.id, 14611063@students.uii.ac.id

***Abstract*—Horticultural commodities are potential commodities that have high economic value and potential to continue to be developed. One type of potential horticultural commodity to be developed is red chili commodity, especially the red chili. The economy of red chili is quite stable in supplying production to the market. In the process of sorting chili by various processing industry companies, chili exporters and farmers who plant red curly chili is generally done manually and involving humans as the chili decision maker is eligible to be elected or not. The process of identifying manually has many disadvantages, some of which are relatively long time required, humans also tend to feel tired and saturated when doing a monotonous activity, differences in perception of quality, product variety is also obtained because of the human visual limitations, as well as strongly influenced by the psychic condition of the observer. The development of science and digital image processing technology makes it possible to sort the agricultural products and plantations automatically. One of Deep Learning methods technique that is Convolutional Neural Network (CNN) method which currently has the most significant result in image recognition is CNN. The classification accuracy value obtained from the training data is 97.14% and the test data is 80% using RGB input image.**

Keywords—chili, deep learning, convolutional neural network

I. INTRODUCTION

Horticultural commodities are potential commodities that have high economic value and potential to continue to be developed. From the supply or production side, Indonesia's vast territory with its agro climate diversity enables the development of various horticultural crops includes 323 types of commodities consist of 60 kinds of fruits commodities, 80 types of vegetable commodities, 66 types of biopharmaceutical commodities and 117 types of ornamental plant commodities.

One type of potential horticultural commodity to be developed is red chili commodity, especially the curly red chili. Some important reasons for the development of this chili commodity is (1) high economic value commodity, (2) the phenomenon of value ladder symptoms of shifting consumer demand from low-value commodities towards high-value commodities horticulture, (3) national and regional flagship commodities, (4) red chili farming is labor intensive, (5) occupy an important position in the menu of food that is consumed every day by almost all the population of Indonesia, (6) consumption of chili by household in the form of fresh chili 80% and for processing industry 20%, (7) a wide adaptation of lowland rice to highland drylands, (9) involving young skilled workforce in rural areas, (10) has considerable diverse benefits and industrial raw materials as well (11) has a variety of market objectives for both traditional markets, modern markets (supermarkets) as well as for the processing industry.

National chili need in one year for all big cities with population of 1 million or more about 800,000 tons per year or 66,000 tons per month. Chili needs on religious holidays and *hajatan* seasons usually increases about 10-20% of normal needs. The level of consumption of red chilies per day in several provinces can be seen in Table I:

TABLE I. LEVEL OF CONSUMPTION OF CHILLIES PER DAY IN SEVERAL PROVINCES

No	Province	Consumption (tons/day)			Total (tons/day)
		Red Chili	Green Chili	Cayenne	
1	DKI Jakarta	42.2	6.8	16.1	65.3
2	West Java	81.0	20.5	97.7	199.2
3	Central Java	55.2	17.1	98.3	170.6
4	Yogyakarta	35.4	2.0	9.7	47.1
5	East Java	30.5	6.2	157.4	194.1

The development of harvested area has a positive growth of 5.99% / year and is quite stable. The development of red chili productivity is experiencing a positive growth of 4.46% / year that the cultivation technology is increasingly cultivated by farmers. The development of red chili production in Indonesia has a tendency to increase by 11.62% / year and is stable. Based on the above description shows that the economy of red chili is quite stable in supplying production to the market.

978-1-5386-5280-0/18/$31.00 ©2018 IEEE

In the process of sorting chili by various processing industry companies, chili exporters and farmers who plant red curly chili is generally done manually and involving humans as decision makers, that chili deserve to be elected or not. The process of identifying manually has many disadvantages, some of which are relatively long time required, humans also tend to feel tired and saturated when doing a monotonous activity, differences in perception of quality, product variety is also obtained because of the human visual limitations, as well as strongly influenced by the psychic condition of the observer. It may also result in inconsistency in the selection process. Manual way is done too much time consuming, so if applied on a large industrial scale required the help of machines on the process.

So to check the quality of chilies, whether or not distributed to traditional markets, supermarkets, industries, or for export purposes, to the end consumed or used for other purposes, the need for a system that can recognize the quality of fruit according to two categories that are feasible or not feasible and do the selection automatically with automated sorting system by utilizing the image of chilies. The system built is expected to be a solution to recognize the feasibility of chili and can be applied.

The development of science and digital image processing technology makes it possible to sort the agricultural products and plantations automatically [1]. Image processing is an alternative to solve the problem. This method has a more sensitive capability because it is equipped with electro-optical sensors that can certainly be more precise and objective than the visual way that is subjective and influenced by the psychic condition of the observer. In the future, image processing is expected to be an option in choosing the eligibility of chili without damaging the sample (object). Using this technology, the physical quality of tomatoes can be determined quickly, inexpensively, with a reliable level of accuracy [2]. There are many methods that can be used to perform image processing, one of them is using Deep Learning (DL) technique that is Convolutional Neural Network (CNN) method which currently has the most significant result in image recognition is CNN. This is because CNN tries to imitate the image recognition system in human cortex visual, so it has the ability to process image information [3]. Using this technology, the physical quality of chilies can be determined quickly, inexpensively, with a reliable level of accuracy. There are many methods that can be used to perform image processing, one of them is using Deep Learning (DL) technique that is Convolutional Neural Network (CNN) method which currently has the most significant result in image recognition is CNN. This is because CNN tries to imitate the image recognition system in human cortex visual, so it has the ability to process image information [3].

Convolutional Neural Network (CNN) is one method of deep learning that can be used for image classification process. CNN has been widely used in many real-world applications, such as Face Recognition Image Classification and Recognition and Object Detection, as it is one of the most efficient methods for extracting important features for non-light tasks. The network on CNN is made with the assumption that the input used is an image. This network has a special layer called the convolution layer, where in this layer an insert image will be processed based on a predetermined filter.

II. LITERATUR REVIEW

A. Convolutional Neural Network

Convolutional Neural Network (CNN) is the development of Multilayer Perceptron (MLP) designed to process two-dimensional data. CNN is included in the type of Deep Neural Network because of the high depth of the network and applied to many image data. In the case of image classification, MLP is less suitable for use because it does not save spatial information from image data and assumes each pixel is an independent feature resulting in poor results [3].

The design of the Convolutional Neural Network is motivated by the discovery of the visual mechanism; the visual cortex presents in the brain. The visual cortex contains many cells that are responsible for detecting light in the small, layered sub-region of the visual field, called the receptive plane. These cells act as local filters above the input space, and the more complex cells have larger receptive fields [4].

In the Convolutional Neural Network (CNN) algorithm applied to the image, the image input will pass through several layers until it is finally classified into one class group. This layer includes the Convolutional layer, Pooling Layer, and Fully Connected Layer (Heaton, 2015). While the learning process is divided into two parts: forward propagation and back propagation.

The convolution layer in CNN performs the functions performed by cells in the visual cortex. In general, the convolutional layer of the convolution layer will detect the edge feature of the image, then the subsampling layer will reduce the dimensions of the feature obtained from the convolution layer, and finally forwarded to the output node through forward propagation process, and the predicted data class is finally determined by Softmax method on dense layer or fully connected layer [4].

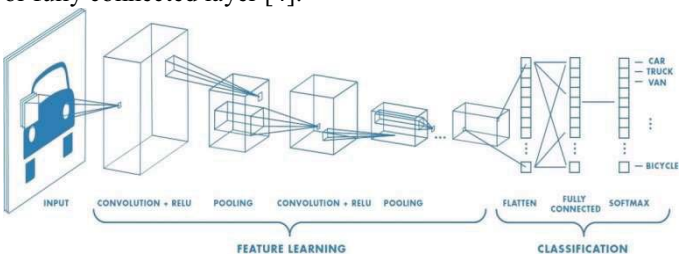

Fig. 1 Convolutional Neural Network

B. Convolution Layer

The first layer to be passed by the input data is the convolution layer. As the name suggests, this layer plays an important role in how CNN works. This layer has several parameters that must be determined by the researcher. Convolution Layer can significantly reduce model complexity through optimizing output. According to the book, some parameters for this optimization are called hyper parameters

consisting of the number of Filters, the size of Filter, Stride, Padding and activation functions [5].

Filter size parameters (f), image size (w), magnitude of shift (s) and padding (p) must be combined in order to make the value of the step count or the value of the step variable being an integer (Heaton, 2015).

$$steps = \frac{w-f+2p}{s} + 1 \quad (1)$$

C. Pooling Layer

In simple terms this layer is responsible for down sampling so that the size of the input data will be smaller. The literature mentions that this layer is not affected by the training phase. This is because this layer has no weight as any other layer [5]. This layer has several parameters that are static, the parameters are spatial extend and stride parameters. Illustration of down sampling on this layer can be seen in Fig. 2 [6].

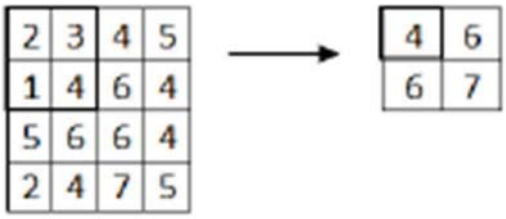

Fig 2. Pooling Layer

The purpose of using pooling or subsampling layer is to reduce the data volume and complexity of the data that will enter in the next layer. The window size used in the subsampling process determines how much information is omitted.

D. Fully-Connected Layer

This layer is a layer that has interconnected neurons. In this layer the feature maps of the previous layer (Pooling layer) will be decomposed into a node. The way this layer works is the same as the forward propagation process in ANN, so in this layer there is also a weight that connects between neurons. These weights will change with the training done on the network [6]. Like neurons on artificial neural networks, neurons on this Dense layer must also have an activation function, in the Deep Learning and Neural Network mentioned that the activation functions that can be used in neurons in this layer are very diverse as ReLU, Sigmoid or hyperbolic activation function [5].

For the number of neurons present in the output layer, the Deep Learning and Neural Network states that the number of neurons used should be the number of classes to be classified (Heaton, 2015). So for example, it will classify the data scattered in 7 classes, then the required output neurons in Fully Connected Layer amount to 7 neurons as illustrated in Fig. 3 [6].

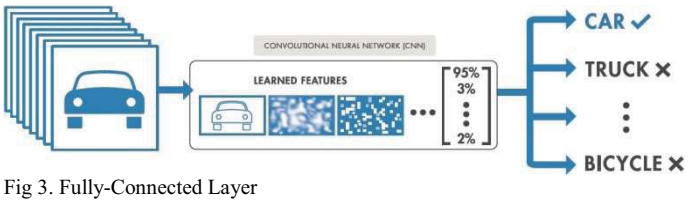

Fig 3. Fully-Connected Layer

III. METHODS

This research uses CNN method for the classification of feasible and unsuitable chilli image. The sample used in this research is 80 images with each category (feasible and not feasible) consist of 40 images.

The type of data used is primary data. The data obtained by taking the image of chili directly by using the camera from smartphone Samsung J5 Prime. The variables used are Worthin Chili and Not Worthin Chili. A worthin chili is the image of red chili that has good quality with fresh red color and no damaged or rotten parts as shown in Fig. 4. While the image of not worthin chilli is the image of red chili pepper with blackish red and there are parts that are damaged or rotten as shown in Fig. 5. Software used is RStudio version 1.1.383. The type of input images is color mode usually called RGB (Red, Green, Blue).

Fig 4. Worthin Chili

Fig 5. Not Wothin Chili

IV. DISCUSSION

A. Model

The model used in this study using the CNN method is a network using 2 convolution layers, 2 layer pooling, 3x3 kernel size, a softmax layer, a fully connected layer, the number of filters on convolution layers 1 and 2 of 32 and 64, after the first

pooling layer is 0.1%, the dropout value after the second pooling layer is 0.01%, the input size is 64 x 64 pixels, the dropout value at the last layer is 0.001%, the number of neurons in the last layer 125 and the data sharing of training and testing by 90%: 10%. The result of the architecture model that can be known how the comparison of loss and accuracy value to the training and testing data shown in the following table.

TABLE II. ACCURACY RESULTS ON DATA TRAINING AND TESTING

Data	Amount of Data	Loss Value	Accuracy Value
Training	70	0.20434	97.14%
Testing	10	0.40000	80.00%

The plot of performance results from the loss and accuracy value generated from the model formed can be shown in the Fig. 6.

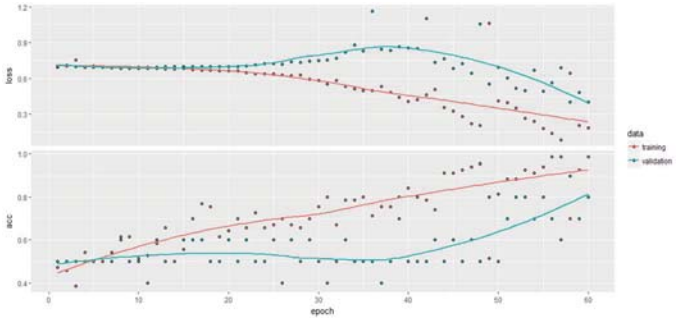

Fig 6. Plot Results of Accuracy and Loss Value

Loss function is a function value to describe the difference between the prediction of the model and the actual data class. A good loss function is a function that results in the lowest error value, while accuracy is a percentage of test data that is correctly classified in the correct class. Based on table 1 it can be seen that the value of loss generated in training data is 0.20434. When compared to the value of loss in the data testing, the value of loss in the training data is smaller so that this value can be said is quite low and good from the model obtained. It is supported by high accuracy value on each data. In the training data has accuration value of 97.14% and 80% in data testing. The resulting model can be said to be able to do a good classification because it has a low loss value and high accuracy value.

B. Classification Results

From the plot image of the performance results of the loss and accuracy value can be seen that the movement of losses are getting close to zero or low as the epoch goes and the movement of accuracy is always increasing this shows good results. The training process will stop if it meets the number of epoch used. In this study the number of epoch that is used as much as 60 epoch, so that the learning process will stop when it comes to the 60th epoch. The results of the classification in training data can be shown in the following confusion matrix table.

TABLE III. RESULT OF CLASSIFICATION IN DATA TRAINING

	Worthin Chilli	Not Worthin Chili
Worthin Chilli	35	2
Not Worthin Chilli	0	33

From the table above can be seen that in the training data there are 2 images classified incorrectly by the model that formed. It can also be seen from the accuracy value generated in the training data that is equal to 97.14%. The classification results of the data testing can be seen in the following confusion matrix table.

TABLE IV. RESULT OF CLASSIFICATION IN DATA TESTING

	Worthin Chilli	Not Worthin Chili
Worthin Chilli	3	0
Not Worthin Chilli	2	5

The classification results in the data testing as shown in table 3 can provide an explanation that not all images are correctly classified into their classes. This can be seen from the results of the classification on the worthin chili. In the table, from 5 images of worthin chili there are 2 classification error. The calculation of the accuracy overall matrix above is as follows:

$$Overall\ Accuracy = \frac{TTP_{all}}{Total\ Number\ of\ Testing\ Entries}$$

$$Overall\ Accuracy = \frac{3}{5} = 80\%$$

So, the accuracy value produced by the model on data testing is 80%.

V. CONCLUSION

Based on the results of the analysis that can be obtained some conclusions. The model used in this study using the CNN method is a network using 2 convolution layers, 2 layer pooling, 3x3 kernel size, a softmax layer, a fully connected layer, the number of filters on convolution layers 1 and 2 of 32 and 64, after the first pooling layer is 0.1%, the dropout value after the second pooling layer is 0.01%, the input size is 64 x 64 pixels, the dropout value at the last layer is 0.001%, the number of neurons in the last layer 125 and the data sharing of training and testing by 90%: 10%. The classification accuracy value obtained from the training data is 97.14% and the test data is 80% using RGB input image.

ACKNOWLEDGMENT

We would like to thank the Directorate of Academic Development, Universitas Islam Indonesia (BPA-UII) for supporting this research financially.

REFERENCES

[1] Kusumaningtyas. Sella and Rosa Andri Asmara, "Identifikasi Kematangan Buah Tomat Berdasarkan Warna Menggunakan Metode Jaringan Syaraf Tiruan", Jurnal Informatika Polinema, ISSN : 2407-070X.

[2] M. Arif Bustomi., "Analisis Distribusi Intensitas RGB Citra Digital untuk Klasifikasi Kualitas Biji Jagung Menggunakan Jaringan Saraf Tiruan," . J. Fisika dan Aplikasinya, vol. 10, October 2014.

[3] Suartika.I. W, Wijaya. A. Y, & Soelaiman, R.. "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101", JURNAL TEKNIK ITS, ppA65-A69.

[4] Hijazi. S, Kumar. R, & Rowen. C, "Using convolutional neural networks for image recognition", Cadence Design Systems Inc.: San Jose, CA, USA,2015.

[5] Heaton. J,"AIFH, Volume 3: Deep Learning and Neural Networks", Washington DC: Heaton Research, Inc,2015.

[6] Priyanta. Ida.B.G.B, "Pengenalan Topeng Bali Menggunakan Algoritma Convolutional Neural Network",Thesis, Fakultas Matematika dan Ilmu Pengetahuan Alam. Yogyakarta :Universitas Gadjah Mada, 2017.