PURPOSE-LED
PUBLISHING™

PAPER • OPEN ACCESS

The Implementation of Deep Learning Using Convolutional Neural Network to Classify Based on Stomata Microscopic Image of *Curcuma* Herbal Plants

To cite this article: U Andayani *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **851** 012035

View the article online for updates and enhancements.

The Implementation of Deep Learning Using Convolutional Neural Network to Classify Based on Stomata Microscopic Image of *Curcuma* Herbal Plants

U Andayani[1], I B Sumantri[2], Andes Pahala[1], **M A Muchtar**[1]

[1]Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia
[2] Department of Pharmacy, Faculty of Pharmacy, Universitas Sumatera Utara,

ulfi.andayani@usu.ac.id

Abstract. There are so many types of herbal plants that come from the same genus. The similarity of features in herbal plants makes it difficult to distinguish. Especially in the pharmaceutical field, which is very risky for making mistakes. The part of plants that is often used as ingredients for herbal medicines is the leaves, so research on leaves and their constituent organelles is very important for the pharmaceutical world. Therefore, an approach is needed to identify the types of organelles present in the leaves, where the organelles most frequently studied are the stomata. So, a neural network approach is needed to distinguish plant characteristics in the same genus. In this research there are 2 species of plants in the *Curcuma* genus, namely turmeric and ginger. This research was conducted by implementing Deep Learning using Convolutional Neural Network (CNN) with the help of the Gabor filter process and feature extraction using the Gray Level Co-occurrence Matrix (GLCM). The study was conducted using 160 microscopic images of Curcuma herbal plants as training data with training accuracy of 93.1% and test data of 40 images with an accuracy of 92.5%.

1.Introduction
The World Health Organization (WHO) defines herbal plants as plants that are used for medicinal purposes and are original ingredients in making herbal medicines [1]. Herbal plants themselves have thousands of species with a total of around 40,000 types of herbal plants that have been known in the world, where around 30,000 species are allegedly located in Indonesia. This amount represents 90% of medicinal plants in Asia. Of these, 25% of them or about 7,500 species are known to have herbal or medicinal properties. However, only 1,200 types of plants have been used as ingredients for herbal medicines.

Based on the Decree of the Head of BPOM Number HK.00.05.4.2411 of 2004 Concerning the Basic Provisions for Grouping and Marking Indonesian Natural Medicines, Indonesian herbal medicines can be grouped into herbs, Standardized Herbal Medicines (OHT) and Fitofarmaka. Jamu only has traditional evidentiary claims, while OHT is a herbal medicine that has been scientifically proven for its quality, safety, and benefits and uses raw materials that have met the standards and are usually pre-clinically tested. While fitofarmaka is a standardized herbal medicine that has been carried out scientifically higher proof, because it has been carried out clinical trials [2].

The various types of herbal plants available, more is used as raw material for herbal medicine, even though it is not curative or healing, but only prevention. Only about 5% of species used as phytopharmaca ingredients, which have been clinically tested in humans can increase curative levels or

can heal. This is caused by one of them due to the lack of further research carried out and the technology base.

One of the most used herbs is from the genus *Curcuma,* with examples of species such as kunyit (*Curcuma longa*) and temulawak (*Curcuma zanthorrhiza*). Both of these species are included in the 10 species of herbs that are most widely used, especially turmeric occupies the second position after ginger of the genus *Zingiber*. The terms genus and species are studied in taxonomy. The taxonomy itself is a grouping of individuals into species, arranging species into larger groups, and naming these groups, so as to produce a classification [2]. The level of classification from the most unique to the level of greater similarity, namely species, genus, family, order, class, division, and kingdom. The similarity of plants is also one of the factors that inhibit the use of herbal plants, which will be very risky if there is an error in choosing the type of plant that is used as medicinal material. The difference can be seen in terms of plant morphology or anatomy.

Plants morphologically or externally have 3 main parts, namely roots, stems and leaves. The leaves are the most important organ for plants in carrying out their lives because plants are *obligate autotrophic* organisms which must supply their own energy needs through the conversion of light energy into chemical energy. The leaves are the most frequently studied parts of plants in biology and pharmacy, especially their microscopic constituents.

Anatomically, the leaves have 3 tissues, namely epidermal tissue (outer layer), mesophyll tissue (inner layer) and transport tissue. In the leaf epidermis tissue, there are stomata. Stomata are easier to study because they are found in the leaf epidermis layer which is the outer layer of leaves, making it possible to be seen directly from a microscope without special treatment such as observation of mesophyll tissue.

Some previous studies on plant leaves have been carried out to resolve identification problems including those carried out by [3] to identify and diagnose diseases in cotton leaves using the *fuzzy* method, *Artificial Neural Network* (ANN), and *Support Vector* (SVM). The researcher also used color and texture extraction in this study. Researchers used 20 diseased leaf images and 25 normal leaf images. Using color segmentation and good image quality will improve diagnosis results. In this study an accuracy of 85% was obtained.

Furthermore, [4] conducted a study to classify the leaves of healthy or affected cocoa plants using *Backpropagation.* In this study an accuracy of 86% was obtained from a total of 90 images data. Subsequent research by [5] conducted a study using *Convolutional Neural Network* (CNN) to detect disease in plant leaves through the detection of artificial neural networks. In addition, researchers also use *invariant moments* as feature extractions. In this study an accuracy of 88% was obtained.

2. Data and Methods Used
2.1.Data Used
The data used in this study were microscopic images of kunyit (*Curcuma longa*) and temulawak (*Curcuma zanthorrhiza*) *plants.* Microscopic data in this study were obtained from research laboratories at the Faculty of Pharmacy, Universitas Sumatera Utara. The image was taken using an electron microscope with a magnification of 10 x 40 magnification scale. As for the training data using 80 microscopic images for each plant, so the total training data there were 160 images, while for the test data using 20 microscopic images for each plant, so the total training data there were 40 images.
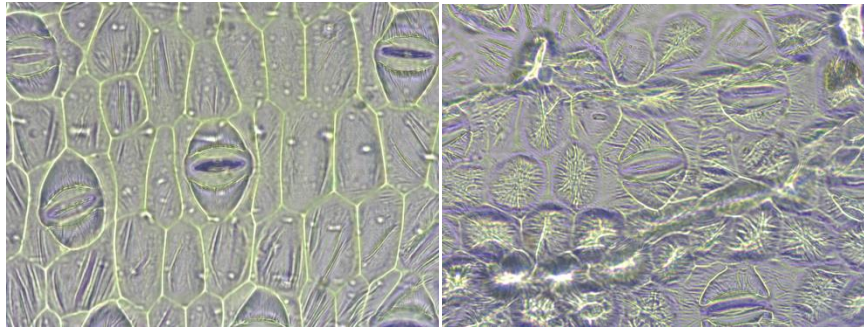
Figure 1. Kunyit stoma (left) and temulawak stoma (right).

2.2. Method Used

The stomata classification system in this study was carried out with several steps, starting from the determination of the dataset, processing, to the resulting output. Here is a general description of architecture.
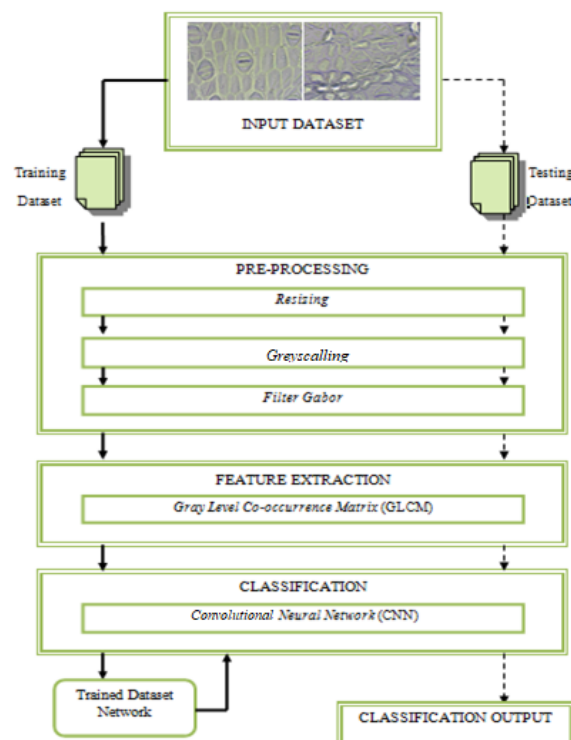

Figure 2. General description of architecture.

2.2.1.Input

Input in the form of stomata microscopic image data from Faculty of Pharmacy, Universitas Sumatera Utara, which was taken using a microscope with a magnification of 10x40.

2.2.2.Preprocessing

Pre-processing stages allow to produce a better image, so that it will be more easily processed and produce a more optimal *output* at later stages.

  1) *Resizing*

  The inputted image will change in size (*resizing*) to 100x100. This process aims to speed up the classification process and generalize the image size of each data.

2) *Greyscale*

At this stage, images that have been resized will be *grayscale* processed into gray images. Since this stomata research require images texture data and not color images, *grayscale* can be done.

Figure 3. Grayscale image.

3) *Gabor Filters*

This stage is done by initializing the following variables:

- Determine the parameters *u, v, m, n* to be used in calculating *the bank filter Gabor*.
  The variable *u* is the number of scales (small to large), with a default value of 5, variable *v* is the number of orientations (horizontal, vertical, diagonal), with a default value of 8. Variable *m* is the number of rows in the 2-dimensional filter Gabor matrix, its value must be odd integer, with the default value is 39, and *n* is the number of columns in the 2-dimensional filter Gabor matrix, the value must be an odd integer, with the default value is 39. Thus, the author sets the following values in the program code:

  gaborFB.u = 6
  gaborFB.v = 9
  gaborFB.m = 39
  gaborFB.n = 39

- Specify the parameters *d1, d2* to be used in calculating *Gabor features*.
  The *d1* and *d2* parameters are *down sampling* of matrix rows and columns. Its function is to reduce the size of the matrix by taking important points from the matrix. The greater the down sampling value, the smaller the size of the resulting matrix, so that calculations can be done faster, but the level of compatibility will be weaker, and vice versa. Following are the values set by the author for the *down sampling* parameter:
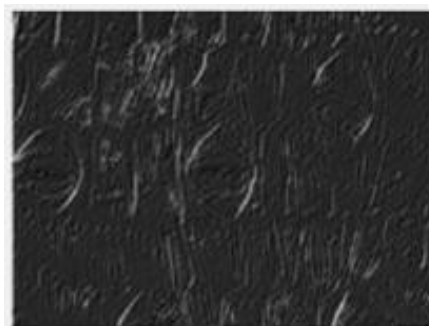 gaborFE.d1 = 3
  gaborFE.d2 = 3

Figure 4. Gabor filter image.

2.2.3. Feature Extraction

Feature extraction is a fundamental part of images analysis. A feature is a characteristic or unique characteristic of an object [6]. Feature extraction used in this study is the *Gray Level Co-Occurrence Matrix* (GLCM). The GLCM feature extraction method is one of the second order extractions of the texture statistical features. Second order extraction shows a statistical relationship between 2 pixels. GLCM is a matrix with the number of rows and columns proportional to the number of *gray levels* (G) in an image.

The GLCM method can produce at least 4 feature extractions from a digital image per pixel neighboring angle. These quantities include *contrast, correlation, energy,* and *homogeneity*.

1) *Contrast*

Contrast is a measure of the existence of gray level variations between image pixels with relative locations. *Contrast* has a limit value from 0 to rank 2 of the symmetric GLCM matrix length. In images with pixel elements that have the same overall value, *contrast* is 0. To find the extraction of *contrast* features, the following formula can be used:

$$Contrast = \sum_{i=1}^{L} \sum_{j=1}^{L} (i-j)^2 GLCM(i,j) \tag{1}$$

2) *Correlation*

Correlation is a measure of linear dependence between gray levels in the image. *Correlation* in GLCM serves to regulate the linear dependence of the gray level in the neighboring pixel image. To find the extraction of features, the following formula can be used:

$$Correlation = \sum_{i=1}^{L} \sum_{j=1}^{L} \frac{(i-\mu_{i\prime})*(i-\mu_{j\prime})*(GLCM(i,j)}{\sigma_{i\prime}*\sigma_{j\prime}} \tag{2}$$

3) *Energy*

Energy is the value of the sum of squares on the elements of the GLCM matrix. *Energy* has a high value when the image has good homogeneity or almost the same pixel value. To find the extraction values, the following formula can be used:

$$Energy \sum_{i=1}^{L} \sum_{j=1}^{L} (GLCM(i,j))^2 \tag{3}$$

4) *Homogeneity*

Used to measure homogeneity, which is a measure of the proximity of the distribution of each element in the GLCM matrix to the diagonal GLCM matrix. To find the extraction of the features, the following formula can be used:

$$Homogeneity \sum_{i=1}^{L} \sum_{j=1}^{L} \frac{GLCM(i,j)}{1+(i-j)^2} \tag{4}$$

2.2.4. Classification with CNN

Convolutional Neural Network (CNN) is a development of the *Multilayer Perceptron* (MLP) which is designed to process two-dimensional data [7][8]. At this stage the images data will be classified into 2 types of output, namely kunyit or temulawak, from the stomata images. Total data used are 300 images that will be divided into training data and test data with a ratio of 250: 50, where 250 training data contains 125 data from each category, while 50 test data contains 25 data from each category.

The initial stage of this image processing is to change the size of the image that was originally sized 100x100 pixels to 32x32 pixels. It is recommended to use images with a power value of 2, such as 16, 32, 64, and so on. For sizes 16x16 and under will make a lot of pixel information disappear, whereas for sizes 64x64 or larger will make the processing rate slower. So the author makes the size value used is 32x32.

In the previous chapter it was explained that the general process on CNN consists of the convolution process, activation function, and pooling. The number of processes in this stage is adjusted to the research needs.

The convolution process is carried out 4 times, or there are 4 *convolution layers* used. In general, already quite two layers are used for the classification model, but in this study used a network

of more *deep* as the implementation of *deep learning* to train the existing models and see how the performance of the model process.

The activation function used is Rectified Linear Unit (*ReLU)* to make the training process faster. The size of the kernel / filter used for each convolution *layer* is 3x3. While the *pooling* size used is 2x2. *Pooling is* done twice, namely after the first two convolution processes and after the next two convolution processes. This is done so that the size of the input is not drastically reduced in each process carried out, so that the input image information that is owned is still useful and can be used in the classification process. While the number of filters / kernels used is also varied, namely as many as 32 filters for the convolution *layers* 1 and 2 as well as 64 filters for the convolution *layers* 3 and 4. The use of more filters in the last two convolution *layers* is because the size of the input at the two layers is smaller, so more filters are needed to extract image information. In the final process of classification, *SoftMax classifier is* used to provide more intuitive results, making it easier to do the classification of probabilistic interpretations for all labels produced.

The process that occurs from the model above is started by "encoding "an image into *features* in the form of numbers that represent that image. In the first convolution process, the image as an *input* size of 32x32 pixels is actually a multidimensional *array* with a size of 32x32x3 (3 is the number of RGB *channels)*. This image will use several processes as mentioned in the model. The filter will be moved by *stride* 1 to all parts of the image input, starting from the upper left corner to the lower right. Every filter shift on the image input is carried out a "dot" operation or mathematical calculation. In this study *zero padding* was not added, so the input size used was equal to 32x32 pixels. In other words, the *output* produced from the convolution process has a smaller size.

The size of the image generated from the first convolution process is 30x30 pixels. The results of this measure can also be obtained using the calculation using equation below:

$$(W-F+2P)/S+1 = 32-3+2(0)/1 + 1 = 30 \tag{5}$$

And so on to do the calculation of the size of the image formed. The results of this operation are then subject to the activation functions, namely ReLU and pooling. The input for the activation function is the real value and the output of the function is a value between 0 and 1. If the input is very negative, the output obtained is 0, whereas if the input is very positive then the output value obtained is 1. In principle, the *pooling layer* consists of a filter with a certain size and *stride* will shift to the entire *feature map* area. *Pooling* used is *Max Pooling.* The purpose of using the *pooling layer* is to reduce the dimensions of the *feature map (down sampling).*

By using a 2x2 size filter and *stride* 2 in the *max pooling* operation, the size of the image formed after the first convolution process is carried out, is subject to an activation function, then the *max pooling* operation is 14x14 pixels. The result of the convolution process is a *feature map* that is reused as *input* for the next convolution process. The process runs until the convolution process ends. In addition, the regularization method used is a *dropout, in* which several neurons will be randomly selected and not used during training. These neurons can be said to be disposed of at random. This means that the contribution of disposed neurons will be temporarily stopped while new tissue and weights are also not applied to the neurons when doing *backpropagation.*

2.2.5 Output
The classification results will display the prediction of plant species in which category based on the image that has been inputted, in this case turmeric or ginger.

3. Research Results and Discussion
This study used training data of 160 images and test data of 40 images of kunyit and temulawak. In this study there are parameters that greatly affect the accuracy of the classification and also the speed of image data processing, such as *Max Epochs* and *Mini Batch.*

Obtaining the accuracy obtained from the testing of the stomata images classification system can be calculated by the percentage of the correct testing data divided by the total amount of testing data:

$$Accuracy = \frac{Total\ True\ Testing\ Data}{Total\ All\ of\ Data} \ x\ 100\%$$

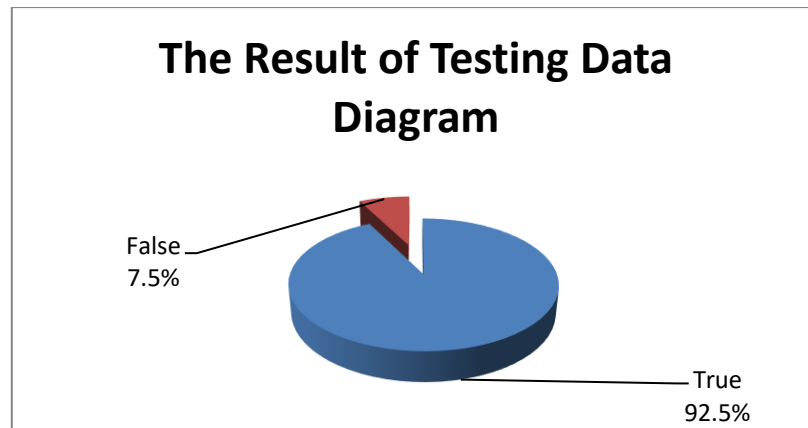$$= \quad 37/40\ x\ 100\%$$

$$= \quad 92.5\%$$

Figure 5. Data result.

From the research conducted, that the accuracy obtained from the training data reached 9 3.1 % accuracy with a total data of 16 0 images, *Max Epochs* 30, *Mini Batch* of 128 and accuracy of test data with an accuracy of 9 2.5 % of the total 4 0 test images shows that the implementation of the *Deep Learning* method using *Convolutional Neural Network* is good enough to classify the microscopic image of herbal stomata from the genus *Curcuma.*

4. Conclusions and Suggestions
4.1. Conclusion
The conclusions that can be drawn from the results of system testing using the *Convolutional Neural Network* (CNN) method for the classification of *Curcuma* plant stomata are as follows:
1. CNN method is able to classify *Curcuma* stomata microscopic image of herbs with an accuracy rate of 93.1% in the training data, and an accuracy of 92.5% in the test data
2. The use of more training data can improve accuracy.
3. Determination of filter size on CNN greatly affects the accuracy results. The smaller the filter size used, the better information obtained.
4. Determination of *Max Epochs* and *Mini Batch* parameters that are too high or too low can affect the accuracy results.

4.2. Suggestions
Suggestions that the author can give for subsequent research are as follows:
1. Establish a broader classification system of herbal plants.
2. Use other classification methods to get higher accuracy.
3. Trying to classify stomata images with other parameters such as stomata type shapes and neighboring cells.
4. Try classification of other microscopic organelles, such as those in stems or roots
.

References

[1]   WHO. 1998. *Guidelines for the Aooripriate use of Herba Medicines*. Manila: WHO Regional Publications, Western Pacific Series No. 23.

[2]   Buletin Infarkes, Kementerian Kesehatan. 2015. *Bugar dengan Jamu*. Direktorat Jenderal Kefarmasian dan Alat Kesehatan, Kementerian Kesehatan. Jakarta.

[3]   Judd, W. S.; Campbell, C. S.; Kellogg, E. A.; Stevens, P. F.; Donoghue, M. J. (2007). *"Taxonomy". Plant Systematics: A Phylogenetic Approach (3rd edition).* Sunderland: Sinauer Associates.

[4]   Permata, E., Suherman, A. & Maulana, A. 2014. *Klasifikasi Daun Tanaman Theobroma Cacao menggunakan metode Neural Network.*

[5]   Gharat, A., Bhatt, K. & Kanase, Bhavesh. 2017. *Leaf Disease Detection Using Image Processing.* The Imperial Journals.

[6]   Sutoyo, T. 2009. *Teori Pengolahan Citra Digital*. Yogyakarta: Andi

[7]   LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. 1998. *Gradient-based learning applied to document recognition*. Proceedings of the IEEE, 86(11), pp. 2278–2324

[8]   Rahmat R F, Dennis, Sitompul O S, Purnamawati S, Budiarto R, 2019 Advertisement billboard detection and geotagging system with inductive transfer learning in deep convolutional neural network *Telkomnika* **17**(5), pp 2659-2666.