Available online at JDR Website: http://journal.unublitar.ac.id/jdr

Journal Of Development Research, 6 (1), May 2022, Pages 1-5

DOI: https://doi.org/10.28926/jdr.v6i1.202

Classification of Toxic Plants on Leaf Patterns Using Gray Level Co-Occurrence Matrix (GLCM) with Neural Network Method

Mohammad Faishol Zuhri(1), S. Kholidah Rahayu Maharani(2), Affandy(3), Aris Nurhindarto(4), Abdul Syukur(5), Moch Arief Soeleman(6)

Universitas Dian Nuswantoro, Indonesia

E-mail: (1)faysme12@gmail.com , (2)kholidahmaharani07@gmail.com, (3)affandy@dsn.dinus.ac.id, (4)arisnurhindarto@gmail.com, (5)abah.syukur01@dsn.dinus.ac.id, (6)arief22208@gmail.com

Received: 22 February 2022; Revised: 8 April 2022; Accepted: 16 April 2022

Abstract

Poisonous plants are plants that must be avoided and not consumed by humans, because the presence of poisonous plants is also often found in the surrounding environment without realizing it. Because of the lack of knowledge to classify poisonous plant species, it will be more difficult to find out. With the help of a computer system, it will be easier to identify the types of poisonous plants. There are 3 types of poisonous plants that will be used in this study, namely cassava, jatropha, and amethyst. There are also 3 types of non-toxic plants with almost the same morphology as a comparison, namely cassava, figs, and eggplant. In this study, researchers tried to classify poisonous plant species using leaf pattern features that would be extracted using shape features and Gray Level Co-occurrence Matrix (GLCM). The value taken from the shape feature is the values of area, width, diameter, perimeter, slender, and round. While the value of contrast, entropy, correlation, energy, and homogeneity for Gray Level Co-occurrence Matrix (GLCM) attributes. To classify data using Neural Network with RapidMiner application. From this study, it is known that from 300 total datasets used, the highest accuracy is 96.13% using the Neural Network method. With an AUC value of 0.986 and is included in the very good category. With this research, it is hoped that it can provide benefits for everyone to know more about the types of poisonous and non-toxic plants with similar morphology in the leaves.

Keywords— Gray Level Co-occurrence Matrix (GLCM), Neural Network, Poisonous plants,

Introduction

Poisonous plants are plants that must be avoided and not for human consumption, because of the importance of plants for humans (Nithaniyal et al., 2021), then the presence of poisonous plants is also often found in the surrounding environment without realizing it (Habibie, 2019). There are many types of plants in the environment that can be consumed by humans and animals, but some plants contain toxins in them (Zheng et al., 2021). From the many types of plants, it makes us less able to distinguish which plants contain toxins and which do not. Because of the types of plants that contain toxins and those that are not, several types of plants tend to have similarities in their morphology (Zheng et al., 2021). So we need a system that can classify the types of poisonous plants.

Along with its development in the world of technology, researchers use models based on the application of image textures, because in this way, computers can recognize the characteristics of a leaf from density, diversity, regularity, and roughness (B et al., 2018). From this process, it can be seen the value of the characteristics or characteristics of the leaf pattern which will be processed for classification. There are several models for analyzing image textures, one of which is by using the Gray Level Co-occurrence Matrix (GLCM) model.

Gray Level Co-occurrence Matrix (GLCM) is a method that is often used as a reference to identify images in texture processing (B et al., 2018). The accuracy obtained from the GLCM method used for feature extraction is by second-order feature extraction (d=2) to calculate the probability of the relationship between pixels that have exactly the same value as the pixel distance (d) and use angles of 0°, 45°, 90 °, and 135°. From the results of extraction with various types of poisonous leaves, for the classification process using the Neural Network method

Materials and Method

I. Material

A. Shape Features

The extraction used in leaf feature extraction uses six types of features, including:

1. Slimness

The slenderness value is taken from the ratio of the length and width of the leaf.

$$Slenderness = \frac{Lp}{Wp} \qquad (1)$$

Where :
lp = leaf length
wp = leaf width.

2. Roundness

The roundness value is taken from the ratio between the area of the object and the square of the perimeter.

$$roundness = \frac{4\pi A}{P^2} \qquad (2)$$

Where :
A = leaf area
P = perimeter of the leaf.

3. Diameter

Diameter is the distance that has the longest size between two edge points of the object.

$$rpd = \frac{P}{D} \qquad (3)$$

4. Perimeter

Perimeter or also known as circumference is the number of pixels contained in the boundary of the shape. Perimeter value is obtained from the edge detection of image objects.

$$Prp = \frac{P}{(L_p + W_p)} \qquad (4)$$

Where :
lp = leaf length
Wp = leaf width
A = area / leaf area.

5. Large

The calculation of the area value is taken from the number of pixels in an object. Here is the calculation to get the area value.

$$Large = \frac{Lp\,Wp}{A} \qquad (5)$$

6. Wide

Is the longest measure of the line that connects the pixels on the edge of the destination that has a perpendicular line to the object with the maximum length.

$$Wide = \frac{D}{Lp} \qquad (6)$$

B. GLCM

This algorithm is used as an assessment of the textures contained in the 2nd order (Neampradit et al., 2018). The measured value taken on the texture of the 1st order uses a statistical assessment based on the pixel value of the original image.This model has at least five attributes in obtaining a value in an image. of the five attributes include ASM, Entropy, IDM, Contrast, Correlation (Ni'mah et al., 2018). There are four angle directions in GLCM, namely the angle of 0°, 45°, 90°, and 135°. The GLCM direction used in this study uses an angle of 90°. (Figure 1 and Figure 2)

1. IDM

Often referred to as a measure of homogeneity which serves to measure the concentration of intensity pairs in the GLCM matrix.

$$IDM = \sum_{i=1}^{L}\sum_{j=1}^{L} \frac{GLCM\,(i,j)}{1+(i-j)^2} \qquad (8)$$

2. Contrast

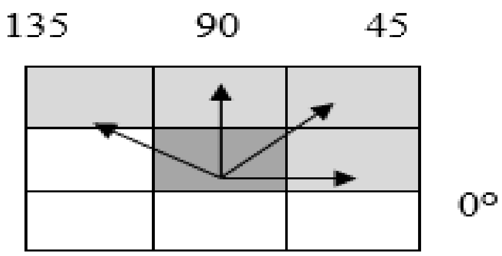


Figure 1. GLCM Angle

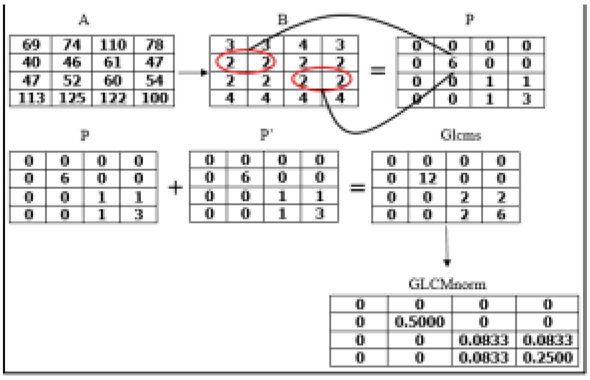


Figure 2. Illustration of the steps of the GLCM method

It is the unequal intensity between one pixel with another pixel with the entire image adjacent to each other. Contrast equation:

$$Kontras = \sum_{i=1}^{L}\sum_{L}^{L}(i-j)^2 GLCM(i,j) \tag{9}$$

3. Energy

Is a component contained in the GLCM which has a function to measure the concentration of intensity pairs in the GLCM matrix. Energy equation:

$$ASM = \sum_{i=1}^{L}\sum_{j=1}^{L}(GLCM(i,j))^2 \tag{10}$$

4. Entropy

It is a measure of the irregularity of the intensity distribution. Entropy equation:

$$Entropy = \sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j) \log GLCM(i,j) \tag{11}$$

5. Correlation

It is the value of the dependence of the gray direction to find the neighboring pixels of the image. the formula can be known as follows:

$$Korelasi = \sum_{i=1}^{L}\sum_{j=1}^{L}\frac{(i,\mu_{i'}) * (j-\mu_{j'}) * GLCM(i,j)}{\sigma_{i'} * \sigma_{j'}} \tag{12}$$

Where :

$$\mu_{i'} = \sum_{i=1}^{L}\sum_{j=1}^{L} i * GLCM(i,j)$$

$$\mu_{j'} = \sum_{i=1}^{L}\sum_{j=1}^{L} j * GLCM(i,j)$$

$$\sigma_{i'}^2 = \sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j)(i-\mu_{i'})^2$$

$$\sigma_{j'}^2 = \sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j)(j-\mu_{j'})^2$$

II. Method

A. Neural Network

Neural Network is a computational model that is almost the same as the structure of a neural network in humans (Liu et al., 2020). *Neural network* belongs to the category of computing science. Neural network describes the ability to resemble the human brain that can perform a certain process, and the output (output) (Ramdhani et al., 2018). Called a neural network because it is designed to follow the workings of the human brain. Pitts and Mc. Culloch, the founder of the neural network method in 1943, showed a mathematical model that could simplify a process derived from real human brain capabilities.

$$y = \int_{i=1}^{n}\left(\sum X_i W_i\right) \tag{13}$$

Information :

Signal X with vector dimension n $(x_1, x_1, \ldots, x_n)$ T will be amplified by synapse w $(w_1, w_2, \ldots, w_n)T$. From the results of the calculations carried out will be changed by the function f. This function can monitor, if the collection of signal gain shows the appropriate excess limit, then the neuron cell which was initially in a state of "0", will produce a signal output of "1". From the output value (y), the neuron has two categories, namely: "0" or "1". And Neurons are said to be firing when they issue an output value of "1". (Figure 3)

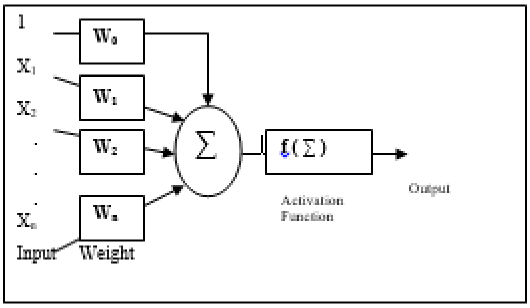
Results and Discussion

For trials in this study, the dataset used for toxic plants in the form of cassava, rubber, amethyst, and castor oil plants. The following is a sample of each poisonous plan. The way the system works is as follows: (Figure 7). In this study, researchers used personal data as data from image extraction results obtained from shape features and the GLCM algorithm, so that the extraction value from the collection of image results could be determined using the neural network method.

A. Data Collecting

The researcher uses images with two categories of plant species, namely poisonous plants and non-toxic plants as comparisons which have almost the same morphology. Where for the category of poisonous plants there are 3 types of plants and for the category of non-toxic plants there are 3 types of plants, each of which uses 50 images for a total of 300 images.

The initial stage after the image retrieval process is to change the white background to remove noise contained in the image and minimize damage during the extraction process. After changing the background color to white, then changing the image size from the original image size of 3120 x 4160 pixels to645 x 879pixels. The following is an example of an image in the early stages of the process until it becomes a gray image.



Information :

1. X_1, X_2.....X_n = signal input
2. W_1, W_2.....W_n = Syneptic link
3. ∑ = Sum
4. f (∑) = Output result

Figure 3. Neuron Model

Figure 4
Cassava Rubber

Figure 5
Amethyst

Figure 6
Distance

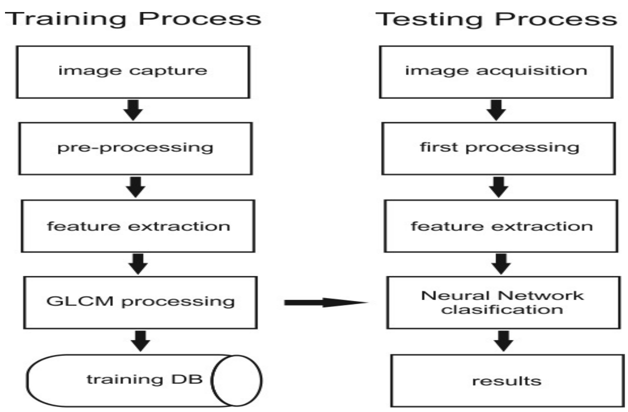


Figure 7. How the System works

B. Data Set

From the shape features and GLCM there are a total of 11 attributes, including: Perimeter, Area, Diameter, Round, Slender, and Width taken from the shape feature. As well as ASM, Energy, Entropy, Correlation and Contrast taken from GLCM.

C. Test

The testing phase is carried out using the RapidMiner application. The result of this test is to know the value of accuracy. The accuracy value obtained from the training data in this trial is 96.19% +/- 4.03% (micro: 96.17%). While the accuracy value of the testing data obtained in the amount of 92.22% with AUC value = 0.989. for accuracy results from class precision and class recall are in (table 1) and (table 2).

Conclusion

This study conducted an image processing experiment to classify toxic plant species with a comparison system using non-toxic plant species that have a morphology of leaf shape that is almost the same as poisonous plants. For the

Figure 8
Original Image

Figure 9
Gray Image

Figure 10
Binary Image

Table 1. Results of Class Precision and Class Recall from training data

	True beracun	True tidak	Class preci-
Pred. Toxic	97	0	100.00%
Pred. N Toxic	8	104	92.86%
Class recall	92.38%	100.00%	

Table 2. Results of Class Precision and Class Recall from training data

	True beracun	True tidak beracun	Class precision
Pred. Toxic	42	4	91.30%
Pred. N Toxic	3	41	93.18%
Class recall	93.33%	91.11%	

application in this research, the researcher uses the shape features and the GLCM algorithm. From the results of trials using the algorithm, shape features have contributed to increasing accuracy with the highest accuracy value of 96.15% using the Neural Network method. With an AUC value of 0.986 and is included in the very good category.

Suggestion

It is hoped that further research can improve the accuracy of comparisons with various methods or by adding other features to extract images.

References

B, M. M., Ana, A., & Hidayat, A. S. (2018). Implementasi Algoritma GLCM Dan MED pada Aplikasi Pendeteksi Kolesterol Melalui Iris Mata. *MIND Journal*, *2*(2), 23–42. https://doi.org/10.26760/mindjournal.v2i2.23-42

Habibie, M. J. (2019). Mengidentifikasi Tanaman Beracun pada Pola Daun dengan Jaringan Syaraf Tiruan Learning Vector Quantification. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, *3*(1), 7. https://doi.org/10.35870/jtik.v3i1.47

Liu, H., Ma, L., Wang, Z., Liu, Y., & Alsaadi, F. E. (2020). An overview of stability analysis and state estimation for memristive neural networks. *Neurocomputing*, *391*, 1–12. https://doi.org/10.1016/j.neucom.2020.01.066

Neampradit, P., Charoenpong, T., Sueaseenak, D., & Sukjamsri, C. (2018). A Method of Thai Main Dish and Soup Classification by Gray Level Co-Occurrence Matrix Algorithm. *IEECON 2018 - 6th International Electrical Engineering Congress*. https://doi.org/10.1109/IEECON.2018.8712294

Ni'mah, F. S., Sutojo, T., & Setiadi, D. R. I. M. (2018). Identification of Herbal Medicinal Plants Based on Leaf Image Using Gray Level Co-occurence Matrix and K-Nearest Neighbor Algorithms. *Jurnal Teknologi Dan Sistem Komputer*, *6*(2), 51–56. https://doi.org/10.14710/jtsiskom.6.2.2018.51-56

Nithaniyal, S., Majumder, S., Umapathy, S., & Parani, M. (2021). Forensic application of DNA barcoding in the identification of commonly occurring poisonous plants. *Journal of Forensic and Legal Medicine*, *78*(September 2020), 102126. https://doi.org/10.1016/j.jflm.2021.102126

Ramdhani, Y., Susanti, S., Adiwisastra, M. F., & Topiq, S. (2018). Penerapan Algoritma Neural Network Untuk Klasifikasi Kardiotokografi. *Jurnal Informatika*, *5* (1), 43–49. https://doi.org/10.31311/ji.v5i1.2832

Zheng, X., An, W., Yao, H., Xu, J., & Chen, S. (2021). Rapid Authentication of the Poisonous Plant Gelsemium elegans by Combining Filter-Paper-Based DNA Extraction and RPA–LFD Detection. *Engineering*, *7*(1), 14–16. https://doi.org/10.1016/j.eng.2020.02.012

