Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 7 No. 2 (2023) 293 - 301 ISSN Media Elektronik: 2580-0760

BPNN Optimization With Genetic Algorithm For Classification of Tobacco Leaves With GLCM Extraction Features

Kristhina Evandari[1], M. Arief Soeleman[2], Ricardus Anggi Pramunendar[3]
[1,2,3]Master of Informatics Engineering, Computer Science, Dian Nuswantoro University
[1]P31202002291@mhs.dinus.ac.id, [2]m.arief.soelema@dsn.dinus.ac.id, [3]ricardus.anggi@dsn.dinus.ac.id

Abstract

Tobacco leaves are one of the agricultural commodities cultivated by Indonesian farmers. In their application in the field, there are many obstacles in tobacco leaf cultivation, one of which is declining tobacco quality caused by weather factors. In this study, a technology-based analysis step was carried out to determine the classification in determining the quality of tobacco leaves. The research was carried out by applying the classification optimization of the Backpropagation Artificial Neural Network Method and genetic algorithms to determine the weights obtained from extracting GLCM features. You can get the weight value from the genetic algorithm on the homogeneity variable from this analysis step. The variable gets a weight value of 1. The results of this study obtained a classification value with the Backpropagation Artificial Neural Network Method model getting an accuracy value of 53.50% at a hidden layer value of 2,4,5,7. For classification with the Artificial Neural Network Method, Backpropagation, which is optimized with genetic algorithms, you get an accuracy value of 64.50% at the 4th hidden layer value. From this study, the value of optimization accuracy increased by 11% after being optimized with genetic algorithms.

Keywords: *tobacco leaves; GLCM; BPNN; genetic algorithms*

1. Introduction

The main problem in this process of growing tobacco certainly depends on climatic factors that farmers cannot control, considering that one of the factors for good production success is to control weather factors[1]. If the weather conditions are erratic and a lot of rainfall falls, tobacco will be too hot and dry long, and then the crop yields will also be dry due to lack of water [2]. This problem certainly makes farmers fail to harvest and get low selling value, so farmers lose money[3]. The determination of weather factors is certainly beyond the control of farmers so; that so far, farmers have only carried out conventional weather calculations following weather climate conditions in Indonesia. Stable weather conditions will help produce good quality tobacco in terms of taste, aroma, and good and superior leaf texture[1]. In this problem, a deeper study will be conducted on the quality of tobacco in the karangawen area, especially in Tlogorejo Village, regarding the quality of tobacco from technology-based science. To find out the superior quality of tobacco in the field of information technology, it is necessary to carry out special and research-based reasoning to obtain the results of new knowledge in the field of information technology[4]. One of the steps that can be done is to apply image-based technology in the branch of informatics engineering to determine the quality of tobacco[2] in Tlogorejo Village. By applying this method in the field of informatics engineering, it is hoped that it will get new analysis results about the quality of tobacco in terms of image texture, and it can be known whether the tobacco has good quality or not. Previous research on tobacco was conducted by Nauval Zabidi Kurniawan[5], who researched the types of leaf diseases of tobacco plants using the Gray Level Co Occurance Matrix texture feature extraction method and the Vector Machine Support Algorithm. This study has been resolved using these methods based on problems arising from tobacco leaves where there are often pests and diseases. So an analysis was carried out to identify tobacco-related diseases and pests based on digital images. From the results of the analysis that have been carried out with the GLCM and SVM methods, testing was carried out with the results of the classification process using the Support Vector Machine algorithm from the GLCM feature extraction value that can be done, the results obtained from the study are 74%. In the use of gaussian (rbf) in the classification, it got the highest success, namely the value of 77% at a distance of 2 pixels. Meanwhile, in

Diterima Redaksi: 05-12-2022 | Selesai Revisi: 13-02-2023 | Diterbitkan Online: 26-03-2023

the polynomial kernel, it was found that the highest success rate was 80% at the distance of the 1st, 2nd, 3rd, and 6th pixels.

The study results above only identify based on diseases and pests in tobacco, even though the quality of tobacco is determined complexly. The main factor influencing the quality of tobacco is the influence of the weather and climate to produce quality tobacco. Therefore, research will be carried out by taking tobacco objects in Tlogorejo with the Backpropagation method based on genetic algorithms and GLCM to identify the quality of tobacco, whether the tobacco has high, medium, or low quality based on weather factors when the tobacco is planted until harvest[6]. To apply the test, research will be carried out using the Back Propagation method using the Gray Level Co-Occurance Matrix texture feature extraction feature.

Backpropagation is a neural network learning algorithm [7]. The field of neural networks was originally ignited by psychologists and neuroscientists who sought to develop and test computational analogs of neurons. A neural network is a set of connected input/output units in which each connection has a weight. During the learning phase, the network learns by adjusting weights to predict the correct class label of the input tuple. BPNN is also referred to as connectionist learning because of the connections between units.

BPNN has attracted many researchers and has emerged as the most popular tool for pattern recognition and classification [8]. This paper presents a sort-based BPNN attribute selection method [7]. This method can determine the relative importance of each attribute, sort the attributes according to their importance, and use BPNN to carry out attribute selection. Since the neural network has inherent shortcomings: based on gradient ancestry, it is easy to fall into the performance of local minimum and weak generalizations. Therefore, another highlight is to use genetic algorithms to optimize the weight and threshold of BPNN, which can significantly reduce the number of iterations and improve the accuracy of convergence to improve the generalization performance [9].

Another research has also been conducted by Taufik Adi Wicaksono, who researched "Analisis Metode GLCM dan SVM untuk medeteksi cacat kain"[10] From the research of this method, which is based on the problem of the absence of a tool or application that can help fabric buyers to carry out the fabric checking process. So that the presence or absence of defects in a fabric can only rely on the accuracy of the eyes of buyers and sellers of fabric, suppose the two do not know the existence of defective fabric. In that case, this is very detrimental and not good for buyers because later, it can be very disruptive and become an obstacle in making and processing fabric into a garment. The results of research conducted by Taufik Adi Wicaksono found that the test results using training tests at the level of system accuracy scores got a score of 83.33%. Meanwhile, the level of test accuracy in the use of supplied test values was 33.33%. Image quality can greatly affect the accuracy of the system that has been built. In addition, the SVM method can also be used to classify defective pattern imagery on fabrics. However, its accuracy still depends largely on the result of extracting the texture features of the Gray Level Co Occurance Matrix method.

As has been done in previous studies, the research will update by applying a Back Propagation artificial neural net based on a genetic algorithm with the GLCM extraction method to analyze the quality of karangawen tobacco [10]. The research researched the Back Propagation method, which raised the classification of processed tobacco leaves. The research introduced the quality of tobacco plant leaves following DMSO (one-processed quality leaves) by analyzing the image of tobacco plant leaves. The first step is to extract the HSV color feature from the results of calculating the HSV value used for the input value in the Back Propagation method. Then based on the trial results, the value of the best learning rate in the value is 0.1, where 19 data can be classified correctly, and 11 cannot be classified correctly. Hence, the resulting accuracy value is 63.33%.

More research on Backpropagation [11] research the prediction of food commodity availability. In this study, the backpropagation method can carry out the process of prediction, but whether or not the value that can be produced will be greatly influenced by the determination of parameters that are the magnitude of the learning rate value, and so can the number of neurons in the hidden layer. In addition, some factors greatly influence the correctness of a prediction in each artificial neural network backpropagation. There is a target error, learning rate, total in the learning data, and the weight value to be given or obtained randomly in each neuron. The results of the test on the architectural model that has and has the smallest RMSE value is the 7-14-1 architecture with an error value of RMSE 0.0033438208, an accuracy value of 99.99% and a performance of 0.2185 so that the trial solution that will be carried out in the study aims to provide that in predicting the amount or total of availability for food commodities Riau Province can provide solutions for the Riau Province Food Crops and Horticulture Office in carrying out the process take a policy on the availability of food commodities in the coming year.

How to apply the backpropagation method to analyze tobacco quality and the classification of predictions of food commodity availability. Therefore a combination of methods will be carried out between the Backpropagation neural network based on the Genetic Algorithm with the extraction feature of the Gray Level

Co Occurance Matrix feature in classifying tobacco quality. This study aims to determine the accuracy level in using the BackPropagation artificial neural network method and optimization of genetic algorithms with texture extraction features, namely the Gray Level Co Occurance Matrix, in classifying the quality of karangawen tobacco.

2. Research Methods

After collecting the data, research will be conducted using good quality and bad tobacco images. Image calculation steps are carried out using the backpropagation method and GLCM extraction features. Here is the flow at the research stage described in figure 1.

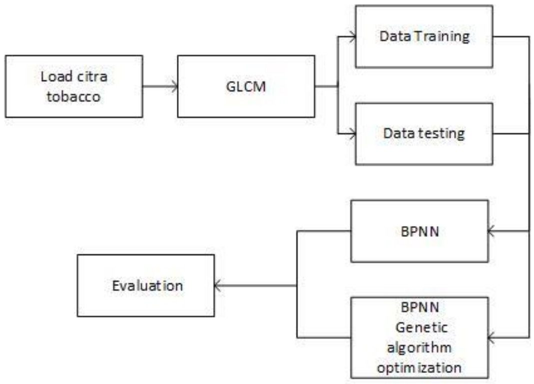

Figure 1 Stages of Research

Details of the training stages of classification data on the quality of karangawen tobacco. The initial stage is to analyze the image or image of tobacco by inputting a folder of tobacco images containing as many as 50 images. Furthermore, the tobacco image is converted to a size of 512 x 512 pixels, and the size is based on one of the standard parameter values present in digital image processing. The values of these standard parameters are intended to keep the digital image circuit simple. The second stage is to pre-process, where the image or images will go from the RGB process to Grayscale in obtaining grayish images[12]. The third stage, thresholding segmentation, is one image segmentation method that separates objects and backgrounds on the image based on the difference in brightness and dark light levels. The fourth stage is extracting texture features using the Gray Level Co-Occurrence algorithm. Extraction of texture features with Gray Level Co-Occurrence. After carrying out the calculation process using the four features of Gray Level Co-Occurrence: contrast, correlation, energy[13], and homogeneity. So that the results in the calculation of the four of these features can be produced from the calculation of the Gray Level Co-Occurrence parameter on the training image, the next stage is to process data with training data and testing data. The sixth stage is the classification process using the Back-propagation algorithm. This stage will be carried out in the image classification process using classes corresponding to those that have been determined. The seventh stage is to carry out the classification process using the Backpropagation algorithm and genetic algorithm optimization to optimize the bias and weight of the Backpropagation algorithm. The seventh stage is the evaluation stage, where this stage will be carried out with calculations with a confusion matrix.

2.1. Dataset

Primary data, data taken from the source directly, and secondary data. The data was a total of 41 images divided into 15 tobacco images of the Temanggung area, 15 of the Kendal area, and 11 of the Karang awen coral area.

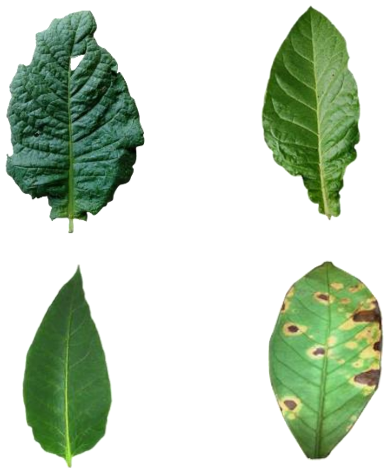

Figure 2. Tobacco leaves

As seen in figure 2 there are differences in the characteristics of tobacco, namely for tobacco Karang Awen has a pointed or conical tip, and the leaves are usually 50 wide and 60 cm long. Temanggung tobacco leaves have an oval-shaped inclined direction shape and are rather wide leafy with a width of 50cm and a length of 60cm. Kendal tobacco leaves have a directional shape leaning oval, 40cm wide, and 50cm long. Somewhat has a smaller tendency than Temanggung tobacco leaves.[2]

2.2. *G*ray Level Co-Occurrence Matrix (GLCM)

The GLCM technique is a way to obtain statistical values of the 2nd order by calculating the odd value by connecting the difference between two-pixel values at a distance with a certain angle [14]. The distance in the GLCM method will create a co-occurrence in the data of an image[6]. In the next stage,[15] you will get one level on the neighboring pixels with pixel values based on distance and orientation in the form of degrees[16]. Pixels can be determined through the distance between pixels which is determined by 1 to 10 pixels, while the angular orientation is formed from four angular

DOI: https://doi.org/10.29207/resti.v7i2.4743

directions, namely 0°, 45°, 90° and 135° [17]. The calculation value sought is as follows[16].

Contrast is the measure of the presence of a value from increasing the grayness value around the image area. There are some differences in the level of color or grayness in the image, and if they have the same neighboring pixel value, then the contrast value is equal to 0. The calculation of contrast can be seen in formula 1.

$$Contrast = \sum_{i,} |i - j|^2 P(i, j) \quad (1)$$

Where i is the matrix row, j is the matrix column, and $P_{(i,\, j)\ is\ the}$ co-occurrence matrix element of rows *(i)* and columns *(j)*.

Correlation is a degree of linear dependence with the degree of grayness in the image that can show a linear value on an image on a neighboring pixel. The value correlates -1 to the value of 1 can be seen in formula 2, 3, 4, 5, and 6.

$$Correlation = \sum_{i,j} = \frac{(i-\mu i)(j-\mu i)P(i,j)}{\sigma i\ \sigma j} \quad (2)$$

$$\mu_i = \sum_i \sum_j ip_{(i,j)} \quad (3)$$

$$\mu_i = \sum_i \sum_j ij_{(i,j)} \quad (4)$$

$$\sigma_i = \sqrt{\sum_i \sum_j (i - \mu_i)}\ p_{(i,j)} \quad (5)$$

$$\sigma_j = \sqrt{\sum_i \sum_j (i - \mu_j)}\ p_{(i,j)} \quad (6)$$

Where i is the matrix row, j is the matrix column, $P_{(i,\, j)}$ is the element of the column co-occurrence matrix of rows *(i)*, and columns *(j)*, μ_i, μ_j is the average of the elements on the rows and columns of the matrix, σ_i, σ_j is the standard deviation on the rows and columns of the matrix sss.

Energy expresses a measure of the irregularity of the grayness in the imagery[18]. Formula 7 shows that the value is high if the GLCM elements have relatively the same value. The low value of elements – GLCM elements with 0 or 1.

$$Energi = \Sigma(i,j)2 \quad (7)$$

Where i is the matrix row, j is the matrix column, and P(i, j) is the column co-occurrence matrix element of row (i) and column (j).

Homogeneity is a measure of the homogeneity (similarity) of an image. The homogeneity value will be of high value if all pixels have a common value.

$$\text{Homogeneity} = \sum_{i_1} \sum_{i_2} \frac{P(i_i i_2)}{1+|i_1+i_2|} \quad (8)$$

Where i is the matrix row, j a is the matrix column, and $P_{(i,j)}$ adalah is the element of the column *(i)* and column *(j)* co-occurrence matrix.

2.3. Back-propagation Neural Network

Backpropagation is a method of learning artificial neural networks that are often used[19]. Backpropagation works using an iterative process using several sample data (training data), then comparing the predicted values on the network of all data examples[6]. In each process, to minimize the Mean Squared Error (MSE) value between the predicted value on the network and the actual value of the relation weight on the modified network[6]. The modification of the neural network relationship is carried out in a reverse direction, from the output layer to the first layer on the hidden layer, so the method is called Backpropagation [7].

As for the error calculation on the output layer, it is obtained using formula 9 [20][21].

$$Errj = (1 - 0i)(Ti - Oi) \quad (9)$$

Oi is the output on the output of the unit node i, Ti is the mean value on the output node in the case example on.

The calculation of errors on hidden layers is done with formula 10 [22].

$$Erri = (1 - 0i)\sum Errjwij\ j \quad (10)$$

Oi is the output of the hidden node of unit I that has output j on the layer, $Errj$ is the value of an error on the unit node j, and WIJ is the weight of both nodes.

Furthermore, after getting the error value on each node, it will be calculated, then modifications to the network weight will be made by using formula 10.

$$wij = wij + l.\ Errj\ .Oi \quad (11)$$

l = is a learning rate with a value between 0 and 1; if l is small in value, then there will be a slight change in weight at each iteration, and vice versa.on The value of the ordinary learning, rate is reduced in the learning process[23].

2.4. Genetic Algorithms

The most popular technique in computational evolutionary research is the algorithmic genetic technique. In traditional genetic algorithms, the representation is a string of bits of fixed length. Each position in the string is assumed to represent a particular feature, and the value stored in that position represents how that feature is expressed in the solution[24]. The string is "evaluated as a collection of structural features of a solution that has little or no interaction." Analogs can be drawn directly to genes in biological organisms [25]. Each gene represents an entity that is structurally independent of another gene. The step of the method of the first genetic algorithm determines the initial population of such data. Furthermore, it determines the fitness evaluation and individual selection. After getting both, Reproduction: Cross Over Mutation will be

DOI: https://doi.org/10.29207/resti.v7i2.4743

carried out. From these results will get the value of the latest population.

3. Results and Discussion

In this study, three tobacco plants were taken from regions already famous for their tobacco products. Namely karangawen tobacco, temanggung and Kendal. In addition, it is mixed with tobacco leaves affected by the disease to identify the type of disease in the tobacco leaves.

3.1. Features of Extraction Gray Level Co-Occurrence Matrix (GLCM)

In using this method, it uses 4 different degree conditions, namely 0, 45, 90.135 (Figure 3). After determining the direction, it further forms the occurrence matrix by calculating the frequency of occurrence pairs of grayness values of reference pixels and neighboring pixels at a specified distance and direction. The calculation uses the orientation value with an angular direction of 0 degrees. 0 degree is the value of images with the same and contiguous values.

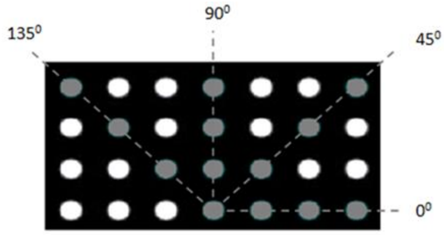

Figure 3 Example Directions for GLCM

In the GLCM calculation, look for 4 variables used as a reference that will be calculated based on the image to be processed. These variables are contrast, correlation energy, and homogeneity. from the results of its implementation, Table 1 data is generated.

Table 1 GLCM Extraction features

No	Contrast	Correlation	Energy	Homogeneity	Label
1	0.154667624755382	0.954942480051130	0.455751183096712	0.938312719137313	Temanggung
2	0.124526051859100	0.968988457053939	0.543916369179699	0.955058122349968	Temanggung
3	0.135510946673190	0.947951129988548	0.474394891633495	0.945658087186073	Temanggung
4	0.100729268590998	0.955238060800546	0.566576761873641	0.959089165953196	Temanggung
5	0.244434931506849	0.952420950770579	0.435074737197162	0.921412199832844	Temanggung
6	0.101348458904110	0.970579918250627	0.481388265784968	0.959216316556588	Kendal
7	0.0739970645792564	0.984435431383591	0.686250625485241	0.974389027336106	Kendal
8	0.0532427226027397	0.985183017573528	0.543327800197302	0.982933662548924	Kendal
9	0.0820236056751468	0.961472492248403	0.700611093684234	0.968766944920091	Kendal
10	0.154288624813784	0.920938323972415	0.589170648763749	0.946135635714222	karangawen
…	………	………	…………	……..	…………..
48	0.167470253254238	0.923968856000874	0.528240345009052	0.937719996987348	karangawen
49	0.207535969914331	0.940822275417087	0.589491138107120	0.941517789326277	karangawen
50	0.132988154916686	0.907861069028923	0.619783006422474	0.950484705047341	karangawen

3.2. Back-propagation Neural Network (BPNN)

Researchers used the Backpropagation Neural Network (BPNN) technique with the results of several degree techniques in the data set [9] extracted with GLCM. In experiments using BPNN, researchers used several hidden layer values to find out which was best related to the value. In this experiment, the hidden layer values are 2,3,4,5,6,7.

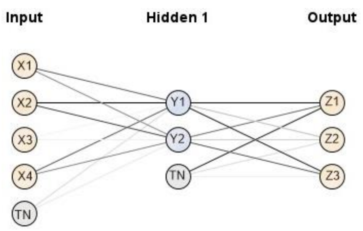

Figure 4 Hidden Layer Node 2

Figure 4 is the use of the 2nd hidden layer. The hidden layer has 2 nodes, for the value of each node is :

For the value of X1 is CONTRAST, X2 is CORRELATION, X3 is ENERGY, X3 is HOMOGENEITY, TN is Bias. for node 1 or Y1 (Sigmoid) the value of CONTRAST is 1.390 and CORRELATION: -2.144, ENERGY: 0.094, HOMOGENEITY: -1.437 for Refractive value: 0.329. On node 2 or Y2 the sigmoid value is CONTRAST: 1.049, CORRELATION: -1.624, ENERGY: 0.130, HOMOGENEITY: -1.073, for Refractive value: 0.191. The output in each class is different. For the Awen Coral class/label on node 1 the sigmoid value is 2,061. For node 2 the sigmoid value is 1.304, and the Threshold value: -1.958. In the Temanggung clas, tobacco leaf class at Node 1: -0.561. For Node 2 value: -0.423 and Threshold value: -0.323. The Kendal tobacco leaf class on node 1 has a Sigmoid value of -1.957, and in node 2 the sigmoid value is -1.456. The threshold in the Kendal tobacco leaf class is 0.126.

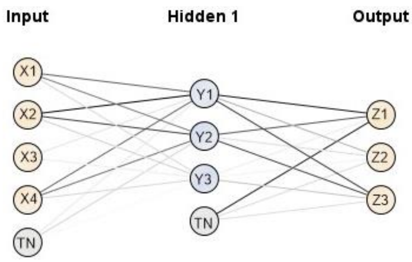

Figure 5 Hidden Layer Node 3

DOI: https://doi.org/10.29207/resti.v7i2.4743

Figure 5 is the use of the 3rd hidden layer. The hidden layer has 3 nodes, for the value of each node is explained below: For the value of X1 is CONTRAST, X2 is CORRELATION, X3 is ENERGY, X3 is HOMOGENEITY, TN is Bias. Node 1 or Y1 values of CONTRAST are 1,340 and CORRELATION: -2,151 ENERGY: 0.218, HOMOGENEITY: -1,300 for Refractive value: 0.175. On node 2 or Y 2 its sigmoid value is CONTRAST: 1.132, CORRELATION: -1.655, ENERGY: 0.018, HOMOGENEITY: -1.239, for Refractive value: 0.355. On node 3 or Y3 its sigmoid value is CONTRAST: 0.329, CORRELATION: 0.397, ENERGY: -0.142, HOMOGENEITY: -0.342, for Refractive value: -0.047.

The output itself in each class is different. For Awen Coral class/label or Z1 Node 1: 1,986 Node 2: 1,458 Node 3: -0.154 Threshold: -1,933 In the Temanggung or Z2 class tobacco leaf class on node 1, the sigmoid value is -0.654, node 2 the sigmoid value is -0.286, node 3 or Z3 the sigmoid value is -0.097 Threshold: -0.306. In the Kendal tobacco leaf class is Node 1: -1,699, Node 2: -1,600, Node 3: -0.498 Threshold: 0.276

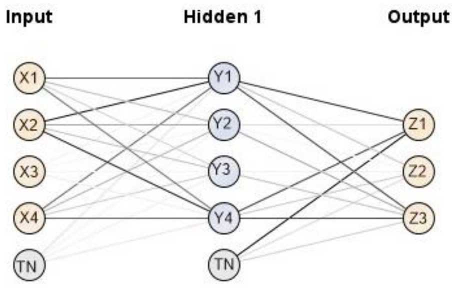

Figure 6. Hidden Layer Node 4

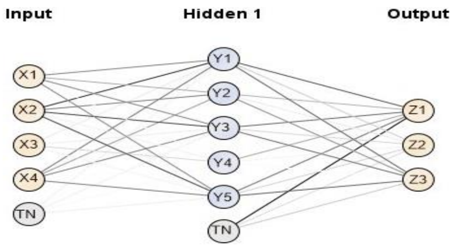

Figure 7 Hidden Layer Node 5

Figure 7 is the use of the 5th hidden layer, the hidden layer has 5 nodes for the value of each node. The value of X1 is CONTRAST, X2 is CORRELATION, X3 is ENERGY, X3 is HOMOGENEITY, TN is Bias. Node 1 or Y1(Sigmoid) CONTRAST: 0.949 CORRELATION: -1.603 ENERGY: 0.151 HOMOGENEITY: -1.047 Bias: 0.102 Node 2 or Y1 (Sigmoid) CONTRAST: 0.640 CORRELATION: -0.983 ENERGY: -0.007 HOMOGENEITY: -0.752 Bias: -0.031. Node 3 or Y1 (Sigmoid) CONTRAST: 0.935 CORRELATION: -1.440 ENERGY: 0.125 HOMOGENEITY: -0.972 Bias: 0.110. Node 4 or Y1 (Sigmoid) CONTRAST: 0.032 CORRELATION: 0.094 ENERGY: -0.295 HOMOGENEITY: -0.004 Bias: -0.046. Node 5 or Y1 (Sigmoid) CONTRAST: 0.935 CORRELATION: -1.433 ENERGY: 0.086 HOMOGENEITY: -1.001 Bias: 0.121.

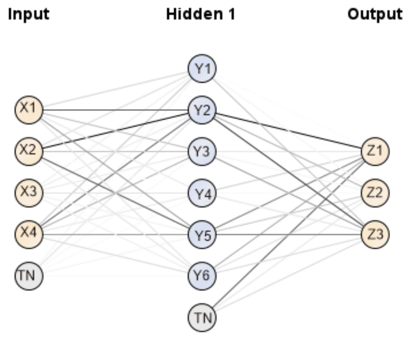

Figure 8 Hidden Layer Node 6

Figure 8 is the use of the 6th hidden layer. The hidden layer has 6 nodes, for the value of each node is explained below. The value of X1 is CONTRAST, X2 is CORRELATION, X3 is ENERGY, X3 is HOMOGENEITY, TN is Bias. Node 1 or Y1 (Sigmoid) CONTRAST: 0.350 CORRELATION: -0.302 ENERGY: -0.130 HOMOGENEITY: -0.478 Bias: -0.075. Node 2 or Y2 (Sigmoid) CONTRAST: 1.402 CORRELATION: -2.206 ENERGY: 0.166 HOMOGENEITY: -1.314 Bias: 0.245. Node 3 or Y3 (Sigmoid) CONTRAST: 0.558 CORRELATION: -0.528 ENERGY: -0.176 HOMOGENEITY: -0.683 Bias: 0.055. Node 4 or Y4 (Sigmoid) CONTRAST: 0.186 CORRELATION: -0.005 ENERGY: -0.237 HOMOGENEITY: -0.189 Bias: -0.043. Node 5 or Y5 (Sigmoid) CONTRAST: 0.852 CORRELATION: -1.483 ENERGY: 0.116 HOMOGENEITY: -0.875 Bias: 0.084. Node 6 or Y6 (Sigmoid) CONTRAST: -0.229 CORRELATION: 0.234 ENERGY: -0.140 HOMOGENEITY: 0.311 Bias: -0.033. Output class 'karangawen' or Z1 (Sigmoid) Node 1: -0.021 Node 2: 2.176 Node 3: 0.220 Node 4: -0.405 Node 5: 1.113 Node 6: -0.738 Threshold: -1.416. Class 'temanggung ' or Z2 (Sigmoid) Node 1: 0.075 Node 2: -0.660 Node 3: 0.090 Node 4: 0.036 Node 5: -0.476 Node 6: -0.262 Threshold: -0.197. Class 'kendal' or Z3 (Sigmoid) Node 1: -0.516 Node 2: -1.714 Node 3: -0.866 Node 4: -0.223 Node 5: -1.146 Node 6: 0.321 Threshold: 0.320.

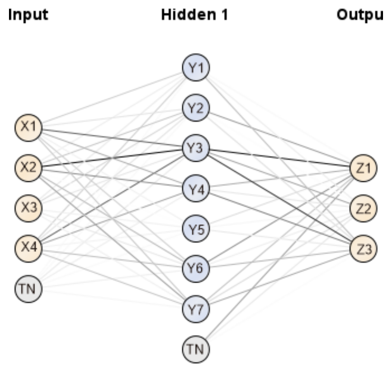

Figure 9 Hidden Layer Node 7

DOI: https://doi.org/10.29207/resti.v7i2.4743

Figure 9 is the use of the 7th hidden layer, the hidden layer has 7 nodes, for the value of each node is explained below. The value of X1 is CONTRAST, X2 is CORRELATION, X3 is ENERGY, X3 is HOMOGENEITY, TN is Bias. Using the 7th hidden layer, for Node 1/Y1 (Sigmoid) CONTRAST: 0.410 CORRELATION: -0.141 ENERGY: -0.182 HOMOGENEITY: -0.413 Bias: -0.042. Node 2/Y2 (Sigmoid) CONTRAST: -0.229 CORRELATION: 0.473 ENERGY: -0.348 HOMOGENEITY: 0.247 Bias: -0.051. Node 3/Y3 (Sigmoid) CONTRAST: 1,357 CORRELATION: -2,155 ENERGY: 0.213 HOMOGENEITY: -1,324 Bias: 0.181. Node 4/Y4 (Sigmoid) CONTRAST: 0.684 CORRELATION: -0.975 ENERGY: 0.002 HOMOGENEITY: -0.774 Bias: 0.143. Node 5/Y5 (Sigmoid) CONTRAST: 0.114 CORRELATION: -0.007 ENERGY: -0.115 HOMOGENEITY: -0.097 Bias: -0.046. Node 6/Y6 (Sigmoid) CONTRAST: -0.411 CORRELATION: 0.602 ENERGY: -0.191 HOMOGENEITY: 0.547 Bias: -0.137. Node 7/ Y7(Sigmoid) CONTRAST: 0.481 CORRELATION: -0.832 ENERGY: 0.128 HOMOGENEITY: -0.491 Bias: -0.071

Output class 'karangawen' Z1 (Sigmoid) Node 1: -0.051 Node 2: -0.843 Node 3: 2.060 Node 4: 0.689 Node 5: -0.288 Node 6: -0.943 Node 7: 0.522 Threshold: -0.822. Class 'temanggung ' Z2 (Sigmoid) Node 1: 0.164 Node 2: -0.013 Node 3: -0.769 Node 4: -0.205 Node 5: -0.066 Node 6: -0.141 Node 7: -0.297 Threshold: -0.206. Class 'Kendal' Z3 (Sigmoid) Node 1: -0.555 Node 2: 0.271 Node 3: -1.744 Node 4: -1.081 Node 5: -0.185 Node 6: 0.607 Node 7: -0.735 Threshold: 0.026

3.3. Genetic Algoritm

In the application of this genetic algorithm, it optimizes variables on the GLCM. This method can determine which variable has the maximum weight in its selection. This study used a genetic algorithm to find the maximum variable to be calculated using the genetic algorithm method. In optimization experiments with genetic algorithms, it uses the Max Number of New Attribute values: 3, Population Size: 5, Maximum Number of Generations: 30, P-initialize, P Crossover, and P Generate of 0.5.

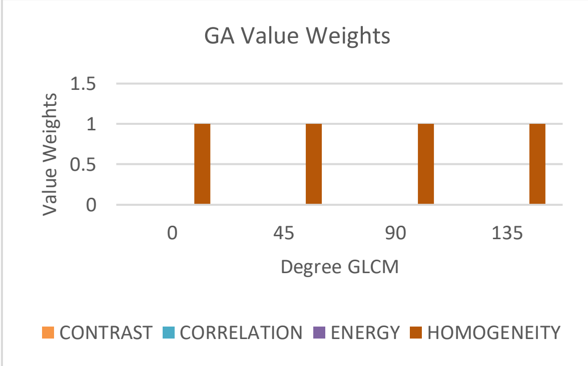

Figure 10 The Greatest Value on Genetic Algorithms

Figure 10 shows the calculation of optimization with genetic algorithms for variable optimization with various extraction feature data with various degrees that have been produced with GLCM, namely 00 450 900 1350 from this optimization that gets a weight with a value of 1 is a homogeneity variable.

3.4. Results of Evaluation and Comparison Accuracy

In the calculation with 10 cross-validations between Backpropagation and Backpropagation, the optimization of genetic algorithms is obtained with a hidden layer 4 value of 64.50%. These results can be seen in the graph.

Figure 11 shows that four hidden layers get a higher accuracy value than other hidden layer values. In the graph, it shows that Backpropagation without a genetic algorithm with three hidden layers gets an accuracy value of 53.50%, and Backpropagation with 4 hidden layers with genetic algorithm optimization gets a value of 64.50%.

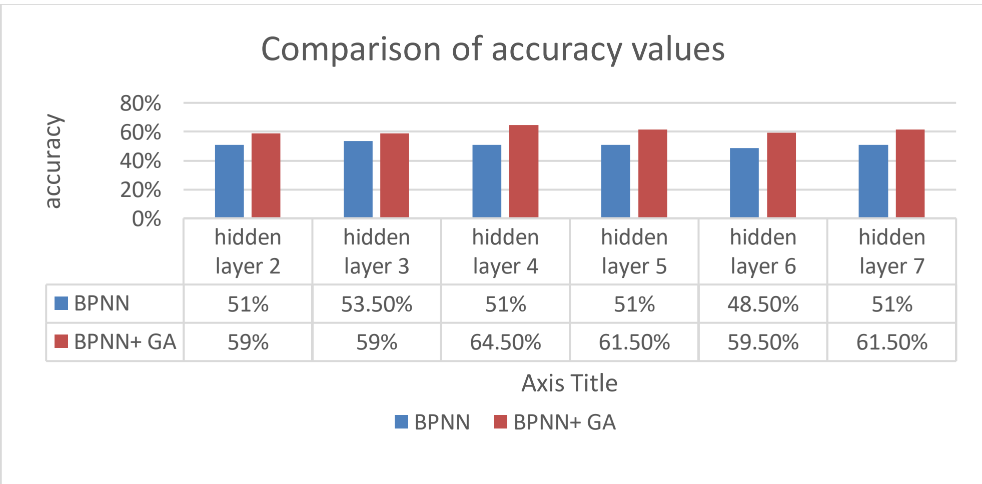

Figure 11 Comparison accuracy

DOI: https://doi.org/10.29207/resti.v7i2.4743

Table 2 Comparison Accuracy

Momentum	learning rate	hidden layer	Acuaration BPNN	Acuaration BPNN+GA	difference
0,9	0,001	2	51%	59%	8%
0,9	0,001	3	53.50%	59%	5%
0,9	0,001	4	51%	64.50%	14%
0,9	0,001	5	51%	61.50%	11%
0,9	0,001	6	48.50%	59.50%	11%
0,9	0,001	7	51%	61.50%	11%
	Average accuracy		51%	61%	10%

4. Conclusion

That genetic algorithm with a maximum attribute value of 1 population size 5 and a maximum generation of 30. In this experiment, the result of an accuracy value of 64.50% on the value of 4 hidden layers. They resulted in an increase in the accuracy value by 10% of BPNN with genetic algorithm optimization. In the results of this study, the increase in accuracy was due to the reduction of the extraction feature variables produced by the GLCM method by the genetic algorithm method, the variables that appeared with values above the average were homogeneity variables. With the accuracy obtained before and after the genetic algorithm, it can be concluded that the genetic algorithm affects the accuracy of Backpropagation, with the most influential attribute being the type of homogeneity

Based on the results of experiments and tests on tobacco leaf classification using GLCM (Gray Level Co Occurance Matrix) for extraction features, as well as classification with Back Propagation Neural Network (BPNN) optimization with genetic algorithms (GA). Concluded has done several experiments with different hidden layer values. The hidden values of this layer are 2,3,4,5,6, and 7. From the best-hidden layer value at 2,4,5,7 hidden layers, get an accuracy value of 53.50%. This value can be obtained with a momentum value of 0.9, a learning rate of 0.01, and a training cycle of 200.

The gray value of homogeneity makes a reference in taking photos in the dataset, therefore it is necessary to pay attention to being able to get a good homogeneity value in using the dataset. Because homogeneity greatly affects the accuracy value in a learning model.

Reference

[1] C. S. Marzan and C. R. Ruiz, “Automated tobacco grading using image processing techniques and a convolutional neural network,” *Int. J. Mach. Learn. Comput.*, vol. 9, no. 6, pp. 807–813, 2019.

[2] Y. Sari and R. A. Pramunendar, “Classification Quality of Tobacco Leaves as Cigarette Raw Material Based on Artificial Neural Networks,” *Int. J. Comput. Trends Technol.*, vol. 50, no. 3, pp. 147–150, 2017.

[3] S. K. Dasari, K. R. Chintada, and M. Patruni, “Flue-cured tobacco leaves classification: A generalized approach using deep convolutional neural networks,” *SpringerBriefs Appl. Sci. Technol.*, no. 9789811066979, pp. 13–21, 2018.

[4] K. Jha, A. Doshi, P. Patel, and M. Shah, “A comprehensive review on automation in agriculture using artificial intelligence,” *Artif. Intell. Agric.*, vol. 2, pp. 1–12, 2019.

[5] P. U. Rakhmawati, Y. M. Pranoto, and E. Setyati, “Klasifikasi Penyakit Daun Kentang Berdasarkan Fitur Tekstur Dan Fitur Warna Menggunakan Support Vector Machine,” *Semin. Nas. Teknol. dan Rekayasa 2018*, pp. 1–8, 2018.

[6] S. Bakheet and A. Al-Hamadi, “Automatic detection of COVID-19 using pruned GLCM-Based texture features and LDCRF classification,” *Comput. Biol. Med.*, vol. 137, no. June, p. 104781, 2021.

[7] C. Zhang, Y. X. Tian, and Z. P. Fan, “Forecasting sales using online review and search engine data: A method based on PCA–DSFOA–BPNN,” *Int. J. Forecast.*, vol. 38, no. 3, pp. 1005–1024, 2022.

[8] P. D. WANANDA, L. NOVAMIZANTI, and R. D. ATMAJA, “Sistem Deteksi Cacat Kayu dengan Metode Deteksi Tepi SUSAN dan Ekstraksi Ciri Statistik,” *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 6, no. 1, p. 140, 2018.

[9] L. Cui, Y. Tao, J. Deng, X. Liu, D. Xu, and G. Tang, “BBO-BPNN and AMPSO-BPNN for multiple-criteria inventory classification,” *Expert Syst. Appl.*, vol. 175, no. November 2020, p. 114842, 2021.

[10] T. Adi Wicaksono and dan John Adler, “Analisis Metode GLCM dan SVM untuk Mendeteksi Cacat Kain,” *J. Tek. Komput. Unikom-Komputika*, vol. 7, no. 1, 2018.

[11] T. Y. Prahudaya and A. Harjoko, “Metode Klasifikasi Mutu Jambu Biji Menggunakan Knn Berdasarkan Fitur Warna Dan Tekstur,” *J. Teknosains*, vol. 6, no. 2, p. 113, 2017.

[12] C. Zhou, K. Yin, Y. Cao, and B. Ahmed, “Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China,” *Eng. Geol.*, vol. 204, pp. 108–120, 2016.

[13] Y. Agussationo, I. Soesanti, and W. Najib, “Klasifikasi Citra X-Ray Diagnosis Tuberkulosis Berbasis Fitur Statistis,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 3, pp. 736–745, 2018.

[14] N. Kaur and V. Devendran, “Novel plant leaf disease detection based on optimize segmentation and law mask feature extraction with SVM classifier,” *Mater. Today Proc.*, no. xxxx, 2020.

[15] S. H. N. Pulung Nurtantio Andono, “Texture Feature Extraction in Grape Image Classification Using K-Nearest Neighbor,” *Resti*, vol. 6, no. 5, pp. 768–775, 2022.

[16] A. Septiarini, Rizqi Saputra, Andi Tejawati, and Masna Wati, “Deteksi Sarung Samarinda Menggunakan Metode Naive Bayes Berbasis Pengolahan Citra,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 5, pp. 927–935, 2021.

[17] A. Rao and S. B. Kulkarni, “A Hybrid Approach for Plant Leaf Disease Detection and Classification Using Digital Image Processing Methods,” *Int. J. Electr. Eng. Educ.*, pp. 1–19, 2020.

[18] S. S. Panna and Betrisandi, “Klasifikasi Kelompok Usia Melalui Citra Wajah Berbasis Image Texture Analysis pada Sistem Automatic Video Filtering,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 3, pp. 429–434, 2019.

[19] Y. Li, Z. Wang, and Y. Wei, “Pathways to progress sustainability: an accurate ecological footprint analysis and prediction for Shandong in China based on integration of STIRPAT model, PLS, and BPNN,” *Environ. Sci. Pollut. Res.*, vol. 28, no. 39, pp. 54695–54718, 2021.

[20] J. Trivedi, Y. Shamnani, and R. Gajjar, “Plant Leaf Disease Detection Using Machine Learning,” *Commun. Comput. Inf. Sci.*, vol. 1214 CCIS, no. September, pp. 267–276, 2020.

DOI: https://doi.org/10.29207/resti.v7i2.4743

[21] A. Reema Matthew, A. Prasad, and P. Babu Anto, "A review on feature extraction techniques for tumor detection and classification from brain MRI," *2017 Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2017*, vol. 2018-Janua, pp. 1766–1771, 2018.

[22] B. Jiang *et al.*, "Evaluating traditional empirical models and BPNN models in monitoring the concentrations of chlorophyll-a and total suspended particulate of eutrophic and turbid waters," *Water (Switzerland)*, vol. 13, no. 5, 2021.

[23] J. Zhao *et al.*, "Prediction of temperature and CO concentration fields based on BPNN in low-temperature coal oxidation," *Thermochim. Acta*, vol. 695, no. May 2020, p. 178820, 2021.

[24] S.N.Sivanandam, *Introduction to Genetic Algorithms*. Springer Berlin Heidelberg, 2008.

[25] S. Katoch, S. S. Chauhan, and V. Kumar, *A review on genetic algorithm: past, present, and future*, vol. 80, no. 5. Multimedia Tools and Applications, 2021.

DOI: https://doi.org/10.29207/resti.v7i2.4743