Vol. 20, No. 1 March 2024 | DOI: 10.33480/pilar.v20i1.5307

IDENTIFICATION OF POTATO LEAF DISEASES USING ARTIFICIAL NEURAL NETWORKS WITH EXTREME LEARNING MACHINE ALGORITHM

Moh. Erkamim[1*]; Ri Sabti Septarini[2]; Mursalim Tonggiroh[3]; Siti Nurhayati[4]

Smart City Information Systems Study Program[1]
Universitas Tunas Pembangunan, Surakarta, Indonesia[1]
https://utp.ac.id/[1]
erkamim@lecture.utp.ac.id[1*]

Informatics Engineering Study Program[2]
Universitas Muhammadiyah Tangerang, Tangerang, Indonesia[2]
https://umt.ac.id/[2]
risabtis@ft-umt.ac.id[2]

Information Systems Study Program[3,4]
Universitas Yapis Papua, Jayapura, Indonesia[3,4]
https://uniyap.ac.id/
mursalim.t@gmail.com[3], nurhayatist.siti21@gmail.com[4]
(*) Corresponding Author

Creation is distributed below Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract—Potato plants have an important role in providing a source of carbohydrates for society. However, potato production is often threatened by various plant diseases, such as leaf disease, which can cause a decrease in yields. Identification of diseases on potato leaves is currently mostly done by farmers manually, so it is not always efficient and accurate. So the aim of this research is to identify diseases on potato leaves with artificial neural networks using the ELM (Extreme Learning Machine) approach and the GLCM (Gray Level Co-Occurrence Matrix) method for feature extraction. The GLCM approach functions to obtain texture features on objects by measuring how often certain pairs of pixel intensities appear together at various distances and directions in the image. Meanwhile, the ELM algorithm is used for image identification by adopting a one-time training method without iteration, which involves randomly determining weights and biases in hidden layers, thus allowing training to be carried out quickly and efficiently. Evaluation of the model by looking for the level of accuracy produces a value of 84.667%. The results show that the model developed is capable of accurate identification.

Keywords: artificial neural networks, ELM, extreme learning machine, GLCM, potato leaf disease.

Abstrak*—Tanaman kentang mempunyai peranan penting dalam menyediakan sumber karbohidrat bagi masyarakat. Namun produksi kentang seringkali terancam oleh berbagai penyakit tanaman, seperti penyakit daun yang dapat menyebabkan penurunan hasil. Identifikasi penyakit pada daun kentang saat ini banyak dilakukan petani secara manual sehingga tidak selalu efisien dan akurat. Sehingga tujuan dari penelitian ini adalah mengidentifikasi penyakit pada daun kentang dengan jaringan syaraf tiruan menggunakan pendekatan ELM (Extreme Learning Machine) dan metode GLCM (Gray Level Co-Occurrence Matrix) untuk ekstraksi fitur. Pendekatan GLCM berfungsi untuk memperoleh fitur tekstur pada objek dengan mengukur seberapa sering pasangan intensitas piksel tertentu muncul bersamaan pada berbagai jarak dan arah pada gambar. Sedangkan algoritma ELM digunakan untuk identifikasi citra dengan mengadopsi metode pelatihan satu kali tanpa iterasi, yang melibatkan penentuan bobot dan bias pada lapisan tersembunyi secara acak, sehingga memungkinkan pelatihan dilakukan dengan cepat dan efisien. Evaluasi model dengan mencari tingkat akurasi menghasilkan nilai sebesar 84,667%. Hasilnya menunjukkan bahwa model yang dikembangkan mampu melakukan identifikasi secara akurat.*

Kata Kunci*: jaringan syaraf tiruan, ELM, extreme learning machine, GLCM, penyakit daun kentang.*

INTRODUCTION

Agriculture has a crucial role in meeting global food needs. One plant that has an important role is the potato plant (*Solanum tuberosum*), which is the main source of carbohydrates for people in various parts of the world. Potato production in Indonesia is increasing gradually from year to year. This increasing market demand for potatoes motivates farmers and producers to increase production to meet market demand. This is proven by potato production in Indonesia in 2022 reaching 1.5 million metric tons, an increase of 10.5% compared to the previous year (Armavillia, 2023). However, potato crop production is often threatened by various diseases, which can cause a reduction in yield and quality. Diseases on potato leaves are one of the main problems faced by farmers. Identification of diseases on potato leaves is currently mostly done manually by farmers, which may not always be efficient and accurate. Usually, identifying diseases on potato leaves is done by looking at the characteristics of the leaves. However, with the large number of potato plants and the large area of land, it is difficult for farmers to identify diseases on potato leaves. Quick and accurate disease identification is a key step in effective crop management. In this context, a digital image processing approach can be applied to solve the problem of identifying diseases on potato leaves.

The image processing process involves a series of techniques to improve quality, extract information, and understand image structure (Rao et al., 2021). Several studies have implemented image processing on plant diseases with leaf images using various algorithms. There is research regarding the identification of Siamese orange plant diseases by applying the K-Nearest Neighbor (KNN) method (Ariesdianto et al., 2021). This research produces an accuracy of 70%, and this approach carries out classification that works based on the distance principle. However, the KNN algorithm is sensitive to outliers in the dataset and can be computationally intensive when using large data. Further research regarding the classification of diseases in rice leaves uses the Support Vector Machine (SVM) approach (Prastyo et al., 2023). The model built produces an accuracy of up to 80% by carrying out classification by searching for optimal hyperplanes that maximize the margin between different classes. However, the SVM approach is sensitive to parameter selection, which can affect performance. The next research on identifying diseases on grape leaves uses the backpropagation artificial neural network approach (Ansah et al., 2022). This approach is able to produce accuracy values of up to 82%. The Backpropagation artificial neural network algorithm is able to identify images through a series of hidden layers; this network produces output that reflects the class or label that corresponds to the image. However, the Backpropagation algorithm experiences convergence problems, causing the training process to potentially slow down or become trapped in a suboptimal local minimum.

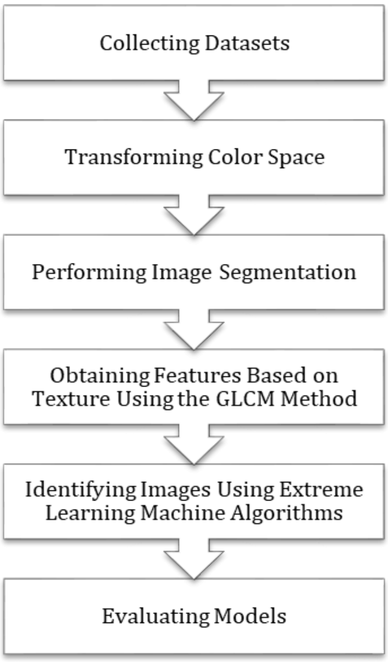
The gap obtained from previous research is that the image identification process requires a fast and effective approach so that it can reduce the computational burden of the process. In the forthcoming study, the identification process will be conducted employing artificial neural networks. This choice is predicated on the recognition of artificial neural networks as an effective method for image identification or classification, owing to their ability to mimic the functionality of human neural networks (Herdiansah et al., 2022). The researchers used artificial neural networks and applied the ELM (Extreme Learning Machine) algorithm to conduct the study. The ELM algorithm is known for its ability to process information quickly and efficiently, so it can be useful in classifying plant diseases with a high level of accuracy (Mayatopani et al., 2021). This fast initialization process allows ELM to skip the iterative and time-consuming training phase typically encountered in traditional neural networks (Wahid et al., 2021). Next, ELM evaluates the model using a closed analytical solution to determine the optimal weights in the output layer without the need for iterative adjustments. For this case study, the leaf characteristics extracted are texture features. The texture feature used is the Gray Level Co-Occurrence Matrix (GLCM), where this approach is used to describe the spatial relationship between image pixel intensities. Extracting features from GLCM provides important information about image structure and texture, enhancing the system's ability to recognize patterns and objects in images (Andrian et al., 2020).

The objective of this study is to discern diseases affecting potato foliage through the utilization of artificial neural networks employing the Extreme Learning Machine (ELM) approach alongside the Gray-Level Co-occurrence Matrix (GLCM) method for feature extraction. It is hoped that the application of this technological framework can lead to the creation of a system that is able to recognize disease quickly and precisely. Therefore, the primary contribution of this investigation lies in the refinement of the ELM algorithm, enabling proficient classification of potato leaf diseases, coupled with furnishing a comprehensive comprehension of these maladies through a meticulous analysis of the traits exhibited by infected foliage.

Vol. 20, No. 1 March 2024 | DOI: 10.33480/pilar.v20i1.5307

MATERIALS AND METHODS

To carry out research, research stages are required, which contain a series of steps or processes carried out in order to carry out research. The stages implemented in this research are illustrated in Figure 1.



Source : (Research Results, 2024)

Figure 1. Research Steps

The stages shown in Figure 1 explain the research steps, which contain the methods used to solve the problem. Further details regarding the research steps are explained as follows:

1. Collecting Datasets

A dataset refers to an organized collection of data, whether in the form of tables, files, or other data structures, used for analysis, research, or machine learning. Datasets in digital image processing refer to collections of image data that are used to train and test algorithms or models for various image processing tasks (Muraina, 2022). The goal of this is to provide enough variation in image conditions so that the model can learn and work well on new data that has never been seen before (Ahmad et al., 2022). The dataset used in this research contains images of potato leaves collected directly using a camera with the same light level. The classes used consist of 3 classes, namely healthy leaves, early blight disease, and late blight disease. A total of 600 images were collected to create the dataset. The data obtained is then distributed into the data used for training and testing with a distribution percentage of 70% and 30%. This division is based on the use of data that is not large, so most of the data is used for training to create more relevant learning patterns (Muraina, 2022). So, we obtained training data for 450 images and testing data for 150 images.

2. Transforming Color Space

Color space transformation refers to the conversion of color values from one color space system to another. This is done to make it easier to represent images and obtain the required information. The color space transformation carried out is from the Red-Green-Blue (RGB) image into the CIE L* a* b* color format, also known as CIELAB. The CIELAB color system was designed to provide color representation that is more consistent and more closely related to human perception of color (Malounas et al., 2024). The CIELAB color space transformation offers several advantages, primarily because it better reflects the way the human eye sees and perceives color (Baek et al., 2022). This color space consists of three main components: L* (brightness), a* (green and red parts), and b* (blue and yellow parts). Additionally, the CIELAB color space transformation offers color universality, where differences in the same color are considered uniform across the color spectrum. The values of L*, a* and b* can be obtained by equations (1), (2) and (3).

$$L^* = 116\left(\frac{Y}{Y_n}\right)\frac{1}{3}16\ for\ \frac{Y}{Y_n} > 0.008856 \qquad (1)$$

$$a^* = 500\left(f\left(\frac{X}{X_n}\right)\right) - f\left(\frac{Y}{Y_n}\right) \qquad (2)$$

$$b^* = 200\left(f\left(\frac{Y}{Y_n}\right)\right) - f\left(\frac{Z}{Z_n}\right) \qquad (3)$$

3. Performing Image Segmentation

Image segmentation is an important step in image processing that deals with dividing or separating an image into different segments based on related properties or attributes (Miao et al., 2023). In this case study, the desired object is separated from the background using the K-Means Clustering approach. This method divides the image into k groups, or clusters, where each pixel is grouped into the cluster that has the average intensity that is closest to the intensity of that pixel (Javidan et al., 2023). To obtain groups through the cluster process in K-Means Clustering, it can be calculated using equation (4).

$$\overline{V_{ij}} = \frac{1}{N}\sum_{k=0}^{n_i} x_{kj} \qquad (4)$$

where $\overline{V_{ij}}$ refers to the center of cluster i in attribute j, ni refers to the number of groups in cluster i, and x_{kj} refers to the value k in attribute j.

4. Obtaining Features Based on Texture Using the GLCM Method

Feature extraction is a procedure that seeks to extract significant data or distinctive attributes from a picture (Borman et al., 2023). The main information in this case study is the texture characteristics of the object, which are represented by the texture feature. One of the approaches used to obtain the texture characteristics of popular objects is the GLCM (Gray Level Co-Occurrence Matrix). The GLCM approach measures the relative appearance of pairs of gray levels that are close to each other in an image (Wanti et al., 2021). This process involves the formation of a matrix that represents the spatial distribution of gray between pixels (Andrian et al., 2020). In each matrix cell, an entry describes the frequency of occurrence of a particular pair of gray values in a particular direction.

Extraction parameters from GLCM, such as contrast, correlation, energy, and homogeneity, provide very useful information for obtaining information about image texture. The contrast parameter is used to measure the contrast between high and low intensity. High contrast values indicate sharp differences between pixel intensities. To produce a contrast value, it can be obtained through equation (5).

$$Contrast = \sum_i \sum_j (i-j)^2 pd(i,j) \qquad (5)$$

The correlation parameter is used to measure the extent to which pixels that have the same or different intensities are correlated in one particular direction. The correlation value can be searched using equation (6).

$$Correlation = \sum_i \sum_j \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \qquad (6)$$

The next parameter is energy, where this parameter functions to measure the homogeneity of the intensity distribution, and the higher the value, the more homogeneous the image. Equation (7) calculates the energy value of an image.

$$Energy = \sum_i \sum_j p_2^d(i,j) \qquad (7)$$

The homogeneity parameter measures the extent to which the intensity distribution approaches a single intensity distribution. A high homogeneity value indicates that the pixel intensity tends to be uniform. Equation (8) calculates the homogeneity value.

$$Homogenity = \sum_i \sum_j \frac{pd(i,j)}{i+|i-j|} \qquad (8)$$

5. Identifying Images Using Extreme Learning Machine Algorithms

ELM is a machine learning algorithm known for its ability to quickly and efficiently train artificial neural networks (Zabala-Blanco et al., 2020). ELM differentiates itself by adopting a one-step training approach where the input weights and hidden weights of hidden layers can be calculated directly and randomly, without the need for iterative adjustments required by some conventional neural network algorithms (Wahid et al., 2021). In the training stage, determining the weighting for the input layer and hidden layers is done with random values and in one step without requiring repeated adjustments.

To prepare the hidden layer in the ELM network through several processes. If N represents the input and the target, which is denoted by $(\boldsymbol{x_i}, \boldsymbol{t_i})$, it refers to $\mathrm{x_i} = [\mathrm{x_{i1}}, \mathrm{x_{i2}}, \cdots \mathrm{x_{in}}]^{\mathrm{T}} \in \mathrm{R^n}$ and $\mathrm{t_i} = [\mathrm{t_{i1}}, \mathrm{t_{i2}}, \cdots \mathrm{t_{in}}]^{\mathrm{T}} \in \mathrm{R^m}$. So, the number of hidden layers is $\tilde{N}$ and the activation function *g(x)* is obtained through equation (9).

$$H \sum_{i=1}^{\tilde{N}} \beta_i g_i(x_{ij}) = \sum_{i=1}^{\tilde{N}} \beta_i g_i(w_i \cdot x_j + b_i) = o_j \qquad (9)$$

where w_i is the notation for the weight vector that connects the input and hidden layers, β_i refers to the weight value connecting the hidden layer and the target, b_i refers to the threshold result in the hidden layer, and $w_i \cdot x_j$ refers to the multiplication value.

From equation (9), simplification is carried out to become equation (10).

$$H\beta = T \qquad (10)$$

where H is the notation for the hidden layer input matrix and T is the notation for the destination matrix.

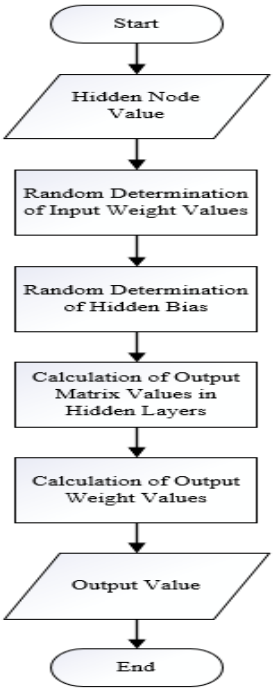
To obtain the input values of hidden weights and biases, random values are given to the network, so that the output weights connected to the hidden layer are found using equation (11).

$$\beta = H \dagger T \qquad (11)$$

where β refers to the output weight value, H refers to the input hidden layer matrix, while T refers to the target matrix.

ELM has a neural network architecture consisting of input and hidden layers and is usually used for classification or identification tasks. In the training stage, the weights between the input and hidden layers are generated randomly and performed in a single step without requiring repeated retuning. For more details regarding the ELM algorithm process, it is visualized in the flowchart in Figure 2.

PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 20, No. 1 March 2024 | DOI: 10.33480/pilar.v20i1.5307



Source : (Research Results, 2024)
Figure 2. Flowchart for the ELM Algorithm Process

The flowchart in Figure 4 illustrates the process of the ELM algorithm, where randomization is used to assign weights to the hidden layers, which contain a greater number of neurons. Next, give weight to the outer layer based on the analysis method with matrix inversion. Below is a description of the stages of the ELM network procedure:

Input : Input pattern (x_i, t_i)
Output : Input weight (w_i), output weight (β_i) and hidden bias (b_i)
The stages :
Stage 1 : Prepare the activation function $(g(x))$, then the hidden node value $(\tilde{N})$
Stage 2 : Get random values for the input weight values (w_i) and hidden bias (b_i)
Stage 3 : Calculate the output matrix value (H) in the hidden layer
Stage 4 : Calculate the output weight value (β)

6. Evaluating Models

The evaluation of this stage measures the performance and accuracy of a model being developed. The evaluation approach relies on the confusion matrix. A confusion matrix is a table that compares the model classification results with the actual values of the data (Wu, 2022). The confusion matrix consists of four main cells: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP is the amount of data that is truly correctly classified into the positive class, while TN is the amount of data that is truly correctly classified into the negative class. FP is the amount of data that should be classified as a negative class but is incorrectly classified as a positive class, while FN is the amount of data that should be classified as a positive class but is incorrectly classified as a negative class. Using these values, several model evaluation metrics can be calculated, such as precision, recall, and accuracy. The study utilized accuracy, precision, and recall evaluations because they offer crucial insights into the performance of the classification system in identifying potato leaf diseases. These values can be obtained through equations (12), (13), and (14).

$$Precision = \frac{TP}{TP+FP} \quad \ldots\ldots(12)$$

$$Recall = \frac{TP}{TP+FN} \quad \ldots\ldots(13)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad \ldots\ldots(14)$$

RESULTS AND DISCUSSION

To identify potato diseases based on leaf images, a dataset is first prepared, which is used as training and testing data. In this research, the dataset used is images of potato leaves collected by taking them directly using a camera with the same light level. The classification involves three classes: healthy leaves, early blight disease, and late blight disease. A total of 600 images were collected, with the data split into 70% for training (450 images) and 30% for testing (150 images). Table 1 displays sample images from each class used in the dataset.

Table 1. Examples of Images Used as Datasets

No	Class Name	Image Sample
1	Healthy Leaves	
2	Early Blight	
3	Late Blight	

Source : (Research Results, 2024)

Table 1 is an example of a dataset used for training and testing data. The developed model is implemented in the MATLAB application. The stages begin with transforming the RGB image into a CIELAB image, to make it easier to represent the image and obtain the necessary information. An example of the results of changing an RGB image to a CIELAB image is visualized in Figure 3.

(a) (b)

Source : (Research Results, 2024)

Figure 3. (a) RGB image; (b) CIELAB Image

Figure 2 (b) illustrates an instance of the outcomes of image transformation conducted in the CIELAB color space. Subsequently, a segmentation process is executed to distinguish images based on similar characteristics, whereby the foreground and background are segregated. In this particular case study, the foreground is isolated from the background utilizing the K-means clustering technique. This approach partitions the image into k groups or clusters, assigning each pixel to the cluster whose average intensity most closely matches that of the pixel. The outcomes of foreground separation employing K-means clustering are depicted in Figure 4.

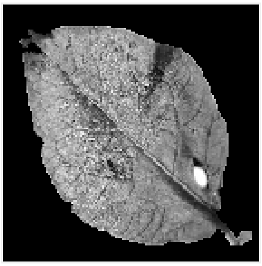
(a) (b)

Source : (Research Results, 2024)

Figure 4. (a) CIELAB Color Space Image; (b) Segmented Image

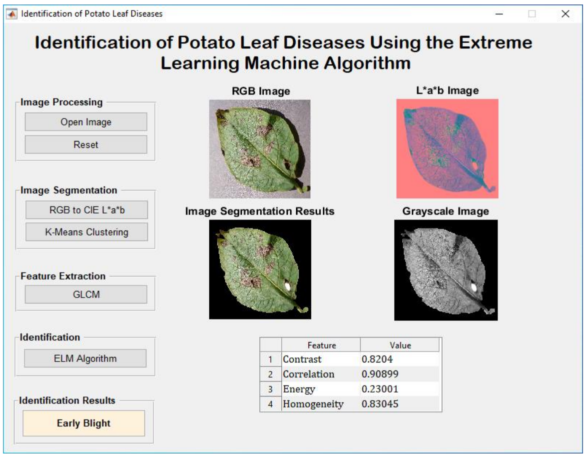
Figure 4 (b) shows the foreground of the segmented image, from which features will be extracted. The main information in this case study is the texture characteristics of the object, which are represented by the texture feature. The texture feature extraction approach applied is the Gray Level Co-Occurrence Matrix (GLCM). In the GLCM approach, it is calculated by measuring how often certain pairs of pixel intensities appear together at various distances and directions in the image. Extraction parameters from GLCM, such as contrast, correlation, energy, and homogeneity, provide very useful information for obtaining information about image texture. Table 2 displays the GLCM parameter values for the segmented images.

Table 2. Results of GLCM Parameter Values

Segmented Image Converted to Grayscale	GLCM Parameters	Value
	Contrast	0.82040
	Correlation	0.90899
	Energy	0.23001
	Homogeneity	0.83045

Source : (Research Results, 2024)

Table 2 shows the input values of all GLCM parameters used for the identification process using the Extreme Learning Machine (ELM) algorithm with artificial neural networks. Next, the model built was tested to measure the performance of the model implemented in MATLAB software. The implementation of the model used for testing in MATLAB software is shown in Figure 5.



Source : (Research Results, 2024)

Figure 5. Model Implementation Interface

To measure the performance of the developed model, an evaluation was carried out using a confusion matrix. The matrix compares the model classification results with the actual data values. Using these values, several model evaluation metrics can be calculated, such as precision, recall, and accuracy. The evaluation process uses 150 test data by comparing model identification findings with actual facts. Figure 6 presents the results of the model evaluation's confusion matrix.

PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 20, No. 1 March 2024 | DOI: 10.33480/pilar.v20i1.5307

Classifier results	Truth data: Healthy Leaves	Early Blight	Late Blight	Classification overall	User's accuracy (Precision)
Healthy Leaves	47	2	2	51	92.157%
Early Blight	3	39	6	48	81.25%
Late Blight	3	7	41	51	80.392%
Truth overall	53	48	49	150	
Producer's accuracy (Recall)	88.679%	81.25%	83.673%		

Source : (Research Results, 2024)
Figure 6. Obtained Confusion Matrix Values

The evaluation resulting from the confusion matrix visualized in Figure 6 obtained measurement values such as precision, recall, and accuracy. The selection of accuracy, precision, and recall assessments in this investigation was deliberate as these three metrics offer crucial and comprehensive insights into the classification system's effectiveness in detecting potato leaf diseases. These values are displayed in Table 3.

Table 3. Precision, Recall and Accuracy Values

Class Name	Precision	Recall	Accuracy
Healthy Leaves	92.157%	88.679%	
Early Blight	81.250%	81.250%	84.667%
Late Blight	80.392%	83.673%	

Source : (Research Results, 2024)

Table 3 displays the outcomes of the model assessment, revealing an overall accuracy of 84.667%. After that, we put these results into criteria groups by referring to the following values: "Very Poor" with a score of less than 40%; "Poor" with a score between 40% and 55%; "Fairly Good" with a score between 56% and 75%; and "Good" with a score between 76% and 100% (Harjanti, 2022). The accuracy results obtained in this study are in the "Good" category.

This research identifies potato leaf diseases by applying an artificial neural network approach using the ELM approach based on texture features using the GLCM method. This research uses traditional machines because they can handle data limitations and have model interpretability. If it is related to previous research which used machine learning to identify plant diseases using leaf images, the results of the proposed model have better results. Research on the identification of Siamese orange plant diseases using the K-Nearest Neighbor (KNN) method produces an accuracy of 70% (Ariesdianto et al., 2021); research on disease classification in rice leaves through the application of the Support Vector Machine (SVM) approach produces an accuracy value of 80% (Prastyo et al., 2023); and research on disease identification on grape leaves uses the Backpropagation artificial neural network approach produces an accuracy rate of 82% (Ansah et al., 2022). Even though the plant object used is different from previous research, namely potato leaves, the proposed model obtains a higher accuracy, namely 84.667%. The reason behind this is because ELM utilizes artificial neural networks, which employ a single training process without any iteration. This process entails the random assignment of weights and biases in the hidden layers. Not only that, the ELM training process involves randomly determining weights and biases in hidden layers, which allows training to be carried out quickly and efficiently. As a result, ELM often provides competitive performance with shorter training times compared to some conventional neural network algorithms.

However, from the model evaluation results, the accuracy value was 84,444%, so the error rate reached 15,556%. This inaccuracy is caused by several factors, including: 1) In the ELM algorithm, the weighting is obtained randomly, which affects the various input values; 2) The ELM algorithm cannot tune parameters adaptively during training, so it is less than optimal for understanding complex relationships in images that require iterative adjustments; 3) The features used are based on texture alone, thereby ignoring other information that might represent the object; 4) The classes used have almost similar characteristics, so to get representative data, pre-processing is needed to get optimal data.

CONCLUSION

Our study created a way to find potato leaf diseases by using artificial neural networks with the ELM (Extreme Learning Machine) algorithm and the GLCM (Gray Level Co-Occurrence Matrix) method to pull out texture features. The GLCM method effectively captures texture features by assessing the frequency of specific pairs of pixel intensities at diverse distances and orientations within the image. These extracted features serve as inputs for the identification process conducted through the ELM algorithm, known for its swift and efficient one-time training approach. Evaluation of the model via the confusion matrix yielded an accuracy of 84.667%, showcasing the model's adeptness in precise disease identification. Nevertheless, enhancements are needed, particularly in refining the weighting process within the ELM algorithm, which currently relies on random assignment. This variability affects input values, suggesting potential integration of complementary algorithms like fuzzy logic to yield more consistent weight values. Additionally, our research underscores the necessity for incorporating diverse feature sets beyond texture alone to enhance the representativeness and robustness of disease identification systems.

REFERENCE

Ahmad, I., Rahmanto, Y., Borman, R. I., Rossi, F., Jusman, Y., & Alexander, A. D. (2022). Identification of Pineapple Disease Based on Image Using Neural Network Self-Organizing Map (SOM) Model. *International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*.

Andrian, R., Maharani, D., Muhammad, M. A., & Junaidi, A. (2020). Butterfly Identification Using Gray Level Co-Occurrence Matrix (GLCM) Extraction Feature and K-Nearest Neighbor (KNN) Classification. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, *6*(1), 11–21. https://doi.org/10.26594/register.v6i1.1602

Ansah, M. A., Kasih, P., & Dara, M. A. D. W. (2022). Identifikasi Penyakit Daun Anggur Berdasarkan Fitur Warna Dan Tekstur Dengan Metode Backpropagation Berbasis Android. *Seminar Nasional Inovasi Teknologi*, 265–271.

Ariesdianto, R. H., Fitri, Z. E., Madjid, A., & Imron, A. M. N. (2021). Identifikasi Penyakit Daun Jeruk Siam Menggunakan K-Nearest Neighbor. *Jurnal Ilmu Komputer Dan Informatika (JIKI)*, *1*(2), 133–140.

Armavillia, K. E. (2023). *Produksi Kentang Di Indonesia Naik Secara Bertahap*. GoodStats. https://data.goodstats.id/statistic/elmaarmavillia/produksi-kentang-di-indonesia-naik-secara-bertahap-P6HbQ

Baek, S.-H., Park, K.-H., Jeon, J.-S., & Kwak, T.-Y. (2022). Using the CIELAB Color System for Soil Color Identification Based on Digital Image Processing. *Journal of The Korean Geotechnical Society*, *38*(5), 61–71.

Borman, R. I., Kurniawan, D. E., Styawati, Ahmad, I., & Alita, D. (2023). Classification of maturity levels of palm fresh fruit bunches using the linear discriminant analysis algorithm. *AIP Conference Proceedings*, *2665*(1), 30023. https://doi.org/10.1063/5.0126513

Harjanti, T. W. (2022). Implementation of Inference Engine with Certainty Factor on Potential Diagnosis of Brain Tumor Disease. *PILAR Nusa Mandiri: Journal of Computing and Information System*, *18*(1), 25–30. https://doi.org/10.33480/pilar.v18i1.xxxx

Herdiansah, A., Borman, R. I., Nurnaningsih, D., Sinlae, A. A. J., & Al Hakim, R. R. (2022). Klasifikasi Citra Daun Herbal Dengan Menggunakan Backpropagation Neural Networks Berdasarkan Ekstraksi Ciri Bentuk. *JURIKOM (Jurnal Riset Komputer)*, *9*(2), 388–395. https://doi.org/10.30865/jurikom.v9i1.3846

Javidan, S. M., Banakar, A., Vakilian, K. A., & Ampatzidis, Y. (2023). Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning. *Smart Agricultural Technology*, *3*, 1–14. https://doi.org/10.1016/j.atech.2022.100081

Malounas, I., Lentzou, D., Xanthopoulos, G., & Fountas, S. (2024). Testing the suitability of automated machine learning, hyperspectral imaging and CIELAB color space for proximal in situ fertilization level classification. *Smart Agricultural Technology*, *8*, 100437. https://doi.org/https://doi.org/10.1016/j.atech.2024.100437

Mayatopani, H., Borman, R. I., Atmojo, W. T., & Arisantoso, A. (2021). Classification of Vehicle Types Using Backpropagation Neural Networks with Metric and Ecentricity Parameters. *Jurnal Riset Informatika*, *4*(1), 65–70. https://doi.org/10.34288/jri.v4i1.293

Miao, Y., Li, S., Wang, L., Li, H., Qiu, R., & Zhang, M. (2023). A single plant segmentation method of maize point cloud based on Euclidean clustering and K-means clustering. *Computers and Electronics in Agriculture*, *210*, 107951. https://doi.org/https://doi.org/10.1016/j.compag.2023.107951

Muraina, I. O. (2022). Ideal Dataset Splitting Ratios in Machine Learning Algorithms: General Concerns for Data Scientists and Data Analysts. *7th International Mardin Artuklu Scientific Researches Conference*, 496–505.

Prastyo, B. A., Istiadi, I., & Rahman, A. Y. (2023). Klasifikasi Penyakit Daun Padi Menggunakan Support Vector Machine Melalui Tekstur Dan Warna Daun Dengan HSV Dan Gabor Filter. *The 6th Conference on Innovation and Application of Science and Technology (CIASTECH)*, 567–575.

Rao, V. C. S., Venkratamulu, S., & Sammulal, P. (2021). *Digital Image Processing and Applications*. Horizon Books (A Division of Ignited Minds Edutech P Ltd).

Wahid, R. R., Anggraeni, F. T., & Nugroho, B. (2021). Brain Tumor Classification with Hybrid Algorithm Convolutional Neural Network-Extreme Learning Machine. *International Journal of Computer, Network Security and Information System*, *3*(1), 29–33. https://doi.org/https://doi.org/10.33005/ijconsist.v3i1.53

Wanti, E. P., Pariyandani, A., Zulkarnain, S., & Idrus, S. (2021). Utilization of SVM Method and GLCM Feature Extraction in Classifying Fish

Vol. 20, No. 1 March 2024 | DOI: 10.33480/pilar.v20i1.5307

Images with Formalin. *Scientific Journal of Informatics*, *8*(1), 168–175. https://doi.org/10.15294/sji.v8i1.26806

Wu, M.-T. (2022). Confusion matrix and minimum cross-entropy metrics based motion recognition system in the classroom. *Scientific Reports*, *12*(1), 3095. https://doi.org/10.1038/s41598-022-07137-z

Zabala-Blanco, D., Mora, M., Hernandez-Garcıa, R., & Barrientos, R. J. (2020). The Extreme Learning Machine Algorithm for Classifying Fingerprints. *39th International Conference of the Chilean Computer Science Society (SCCC)*. https://doi.org/https://doi.org/10.1109/SCCC51225.2020.9281232