九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Texture-Based Classification of Benign and Malignant Mammography Images using Weka Machine Learning: An Optimal Approach

Sumarti, Heni
Department of Physics, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang

Sheilla Rully Anggita
Department of Physics, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang

Hartono
Department of Physics, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang

Fachrizal Rian Pratama
Department of Physics, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang

他

https://doi.org/10.5109/7151705

出版情報：Evergreen. 10 (3), pp.1570-1580, 2023-09. Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Japan
バージョン：
権利関係：Creative Commons Attribution-NonCommercial 4.0 International

KYUSHU UNIVERSITY

Texture-Based Classification of Benign and Malignant Mammography Images using Weka Machine Learning: An Optimal Approach

Heni Sumarti[1,*], Sheilla Rully Anggita[1], Hartono[1], Fachrizal Rian Pratama[1], and Alvania Nabila Tasyakuranti[1]

[1]Department of Physics, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang, Indonesia

*Author to whom correspondence should be addressed:
E-mail: heni_sumarti@walisongo.ac.id

(Received November 30, 2022; Revised May 2, 2023; accepted July 31, 2023).

Abstract: Breast cancer is the most common cancer in Indonesia. One way to detect it early is by screening using mammography. Previous trials showed that mammography screening in women aged 40-49 years could reduce breast cancer mortality by 25%. However, misdiagnosis may occur abaout breast density and the patient's physical size due to machines. In addition, human reader errors can occur concerning the reader's experience and perception. Therefore, diagnostic aids are needed to distinguish benign and malignant cases and receive appropriate treatment. The methodology in this research consists of three stages: preprocessing, texture feature extraction, and data classification. Preprocessing consists of filtering, contrast, cropping, and resizing, while texture feature extraction consist of Histogram and (Gray Level Co-occurrence Matrix) GLCM. Data classification using Support Vector Machines (SVM), Naive Bayes, Multi-Layer Perceptron (MLP), Multiclass classifier, and Random Forest methods with Weka Machine Learning software. It produces an accuracy of 62.00%, 62.00%, 88.00%, 82.00%, and 100.00%, respectively. The results of data classification using the Random Forest method show that the accuracy, specifications, and specificity reach 100%. Random forest can be used as the most optimal classification method to distinguish benign and malignant cases based on texture features in mammography images using Weka Machine Learning software. This can help radiologists and medical professionals to diagnose cases and take further steps, such as therapy.

Keywords: Image Processing, Texture Feature, Classification Method, Weka Machine Learning

1. Introduction

Breast cancer is the most common cancer in Indonesia[1)]. Breast cancer is caused by the uncontrolled growth of abnormal cells in the breast and it is thought that these cells have spread from the breast to other tissues, lymph nodes or other areas of the body. It is essential to find these unwanted cells and stop their multiplication as soon as possible to prevent later stages effects. Malignant cells can become malignant and spread to other body parts, while benign cells do not become cancer. If a tumour is detected, the first thing the doctor will do is determine whether the tumour is benign or malignant. This is caused by the different treatment and prevention strategies for these two types of cancer. The difficulty with all these diseases is that there are no screening tools of similar quality that can detect cancer in its early stages. If such a device existed, the patient would be ready to start treatment as soon as possible and work to stop unwanted cell growth or malignancy. However, currently, there is no such machine[2)].

Breast Self Examination (BSE) is the easiest and most cost-effective method for diagnosing breast cancer early. This method uses your eyes and hands to find changes in your breasts. This examination can be done routinely at home without needing any tools. Meanwhile, diagnostic techniques for breast cancer include biopsies, breast ultrasound, mammography, and MRI. Digital mammography is the most common, affordable, and effective breast cancer screening method. This method can detect up to 90% of breast cancer cases even before a lump can be felt, using low-dose x-rays to produce images of the breast in which tissue, including tumors, appears distinctly gray in the image. Therefore, in less developed regions and middle-income countries, digital mammography is the primary choice in diagnosing breast tumour[3)].

Previous trials within ten years (1990-1997) showed

that screening using mammography in women aged 40-49 years could reduce breast cancer mortality by 25%[4]. Radiologists diagnose breast cancer using mammograms and look for signs of abnormalities such as distortion of the breast tissue architecture, differences in alignment between the two breasts, presence of masses, and calcifications[3]. However, misdiagnosis may occur of breast density and physical size of the patient and due to machines, where some lesions are challenging to visualize. In addition, human reader errors can occur to the reader's experience and perceptual and cognitive scrutiny[5]. Therefore, diagnostic tools are needed to distinguish benign and malignant cases and receive appropriate treatment. This has caused many researchers to develop artificial intelligence to facilitate medical experts, one of them is machine learning.

Machine learning is a technology that enables machines to learn independently without required guidance from humans. This technology is based on statistical, mathematical, and data mining principles that allow machines to analyze and learn about that data without human intervention or programming changes. One example of software that can perform machine learning tasks is Weka, which can run various algorithms for information retrieval and data mining processes. Weka has various algorithms that support the object classification process and make it easier for users to implement it directly. Users can load datasets, choose classification algorithms, and are given several data representations that represent the results of the accuracy and error rate of the classification process[6]. The dataset that can be used in medical diagnosis is a series of numbers extracted from digital images. The extraction commonly used in differentiating benign and malignant cases is the extraction of texture features.

Texture features are an image's condition, size, and pixel placement. We can extract texture features from all images, so this method is very effective for image processing. Texture features are characteristics of an image whose features can be used to identify a particular area of an image[7]. Texture features can be adapted to mammography density assessment. The combination of density and texture may improve screening sensitivity in mammography images[8]. Two types of statistical features that can be used to distinguish between different object textures are first and second-order features. First-order features are obtained from the characteristics of the image histogram and are often used to distinguish macrostructural textures. For example, local patterns repeat periodically. Meanwhile, second-order features are based on the probability relationship between two pixels at a certain distance and angular orientation and are used to distinguish microstructural textures such as local patterns and less obvious repeats[9].

The previous review by Fatima et al.[10] discussed various machine learning methods, deep learning, and data mining related to breast cancer prediction. It reviewed 27 machine learning publications, 4 articles on related challenges, and 8 on convolutional neural networks in breast cancer research reviews. The Author observes that most publications use images, but some articles use genetics in their research. Support vector machines (SVM), decision trees, and random forests are the main algorithms used in the genetic analysis of breast cancer. On the other, imaging methods use various algorithms, such as CNN and Naive Bayes. Meanwhile, Pang et al.[11] reviewed the current work applying deep learning to breast cancer using different imaging modalities. The investigations focus on data sets, architectures, applications, and related assessments. It focuses on developing a deep learning framework for mammography using three different modalities (ultrasound, mammography and MRI). They wanted to provide the latest results in breast cancer imaging by implementing a DLR-based CAD system, and this was the focus of their efforts. The study combines the use of classified datasets and CNNs for classification.

Previous researchers have developed various methods of classifying mammographic images to distinguish benign and malignant cases. Previous studies used GLCM and SVM to distinguish benign and malignant cases based on mammography images. It aims to find the optimal GLCM angle for mammogram classification of tumour cases. The mammogram are provided by the Digital Database Screening Mammography (DDSM) data set. According to the experimental results, the accuracy is 63.03% with a specificity of 89.01%[12]. Another study aims to review the performance of artificial intelligent in Multilayer Perceptron (MLP), Neural Networks, and Convolutional Neural Networks (CNNs). It is used to detect breast melanoma for early breast cancer diagnosis. It made an in-depth comparison of the function and design of each tissue. Then an analysis was carried out based on the diagnostic tests and tissue-based classification of breast malignancy to decide which is better. CNN provided slightly better accuracy than MLP for breast cancer diagnosis and detection[13].

Previous research aimed to identify tumours in the breast and differentiate benign from malignant animals using machine learning, a subfield of artificial intelligence. This test uses a Contrast Limited Adaptive Histogram (CLAHE) to show suspicious areas. Create a new data set with feature selection using LBP (Local Binary Pattern), HOG (Histogram of Oriented Gradients), Correlation (COR), and GLCM for feature extraction. Furthermore, features were classified into three categories (normal, benign, and malignant) and two categories (normal and abnormal). We apply different machine learning algorithms (CART, Naive Bayes, C5.0, and Random Forest) to compare their performance. Based on study results to determine whether or not breast tumours produce an accuracy of 65.70% with the Naive Bayes algorithm[14]. Meanwhile, another study compared the effectiveness of the proposed features on individually

annotated data sets obtained from Superspeciality Cancer Hospital, New Delhi, and public data sets, including digital databases for mammography screening, INbreast database, and the Mammographic Image Analysis Society database. This study aims to validate the effectiveness of the Hanman Hesitancy-based transformation classification on mammographic images. The proposed method limits the "non-data-intensive" aspect classification on public and private datasets by providing 100% accuracy for multiclass, outperforming modern methods available in the industry[15]. Previous studies using subjects from the Digital Database for Screening Mammography (DDSM) comprised of 966 images (322 normal, 322 benign, and 322 malignant). This research presented an automated and computer-based approach to classify breast microcalcifications on mammography images using discrete wavelet transform–random forest (DWT-RF). The results show an accuracy of 95%, sensitivity of 93%, and specificity of 97%[16].

Feature extraction and data classification methods to distinguish benign and malignant cases in mammography images have been carried out as described above. Recent findings show that classification using the random forest to distinguish mammogram calcifications cases using Weka machine learning delivers outstanding performance[17]. However, no one has compared classification data based on texture features to distinguish benign and malignant cases using Weka machine learning software. The results of previous studies show that the complexity of information affects engine performance[18]. This is very helpful in differentiating various cases to facilitate diagnosis[19]. Weka machine learning software is open-source software that is very easy to use, so it is user-friendly if used for health workers. In addition, this software is more accurate in data classification than Rapidminer[20]. This study aims to obtain an optimal classification to differentiate between benign and malignant cases with the highest accuracy. Classification was done using texture feature-based Weka machine learning software to diagnose breast cancer cases on mammogram images. It is hoped that the results of this study can assist radiologists in making early diagnoses, thereby improving patient prognosis and reducing mortality from breast cancer. In addition, the use of tools that are practical, affordable, and easy to use will have a positive impact on the economic and social sciences by facilitating public access to technology in the health sector.

2. Methods

This study using mammography digital images from free database source, the Digital Database for Screening Mammography (DDSM). DDSM is a resource used by the Mammography Image Analysis Research Community. Primary support for this project was a grant from the US Breast Cancer Research Program. Army Medical Research and Materials Command[21]. The mammographic images used in this study consisted of 25 benign and malignant cases, respectively. The image taken is grayscale, and the form is displayed as a gif with a marker indicating the location of the tumour mass.

The procedure used in this study consists of three stages, as shown in figure 1. It consists of three stages: preprocessing, texture feature extraction using MATLAB version R2013b, and data classification using Weka Machine Learning software. Furthermore, the data is processed using the confusion matrix generated from the classification process by Weka Machine Learning software and compared to the results of the classification method that provides the highest accuracy.

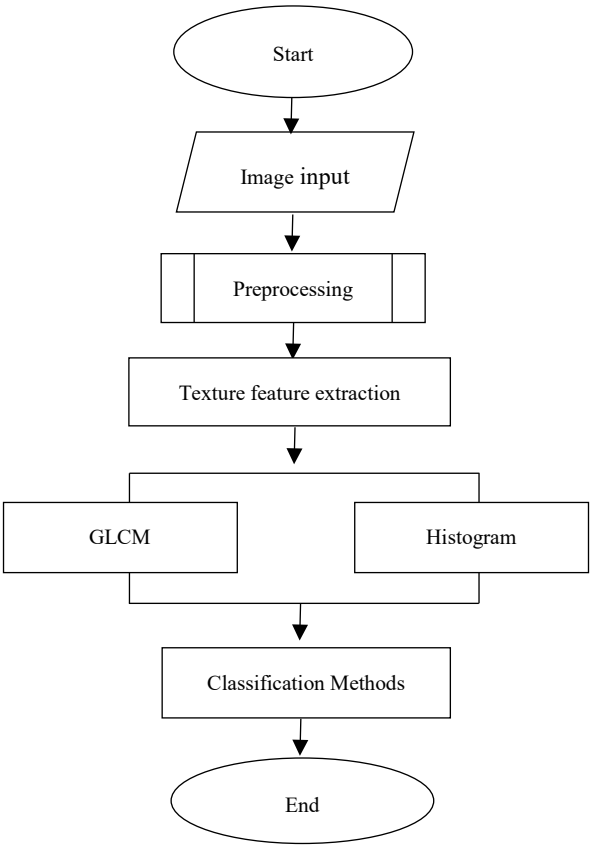


Fig. 1: Research procedure

2.1 Preprocessing Image

Preprocessing images includes filtering, contrast, cropping, and resizing mammographic images. The filter used in this study is the median filter. Image filtering is performed to remove the marker in the tumour in mass the mammography image. Contrast enhancement using histogram equalization. Image contrast aims to improve image quality to clarify the difference between benign and malignant cases in mammography images. Cropping is done on the part of the tumour mass that is being processed. Resizing is intended to even out the size of the mammography image.

2.2 Feature Texture

We performed a texture feature extraction to get the statistical data from the mammography image, which has

been preprocessed. This study uses a texture feature extraction process with two statistical orders, a histogram and a Gray-Level Co-occurrence Matrix (GLCM). The attributes used in the histogram are mean, standard deviation, variance, entropy, skewness, and kurtosis. Meanwhile, the attributes of GLCM are contrast, correlation, energy, and homogeneity.

2.2.1 Histogram

The histogram describes the frequency of each intensity value that appears throughout the pixels during image processing. It is useful in observing the sadness of intensity values intensity[22]. This indicator was used as a basis for adjusting the contrast and brightness of an image. Here is the histogram of the following texture features that have been used in this research[23]:

2.2.1.1 Mean

Mean is the average of the data in pixel images.

$$mean = \sum_{i=0}^{L-1} i.p(i) \quad (1)$$

where: i is the grey level in images, p(i) is the probability of appearing in i, and L is the highest grey level.

2.2.1.2 Standard Deviation

Standard deviation (σ) shows the distribution of data in the image.

$$\sigma = \sqrt{\sum_{i=1}^{L-1}(i - mean)^2 p(i)} \quad (2)$$

2.2.1.3 Variance

Variance is the probability distribution of the image pixel data.

$$\sigma^2 = \sum_{i=1}^{L-1}(i - mean)^2 p(i) \quad (3)$$

2.2.1.4 Entropy

Entropy is a measure of the degree of disorder.

$$Entropi = -\sum_{i=0}^{L-1} p(i) log_2(p(i)) \quad (4)$$

2.2.1.5. Skewness

Skewness is data distribution skewed to the left, right or symmetrical.

$$Skewness = \sum_{i=1}^{L=1}(1 - mean)^3 p(i) \quad (5)$$

2.2.1.6 Kurtosis

Kurtosis is a value that shows data distribution tends to be flat or pointed.

$$Kurtosis = \sum_{i=1}^{L-1}(1 - mean)^4 p(i) - 3 \quad (6)$$

2.2.2 Gray-Level Co-occurrence (GLCM)

Gray Level Co-occurrence Matrix (GLCM) is a method for extracting surface texture features of images by shifting the matrix in a certain distance and direction. GLCM is a grid that describes the repetition of occurrences of two matrix pixel intensities with a certain distance and direction in the image[24]. In this research, we use distance one and direction 0°. Also four main texture features[25]:

2.2.2.1 Contrast

Contrast is the difference in gradation, brightness, or tone (colour) between dark and light areas.

$$Contrast = \sum_{n=0}^{N_{g-1}} n^2 \left\{ \sum_{i=1} \sum_{j=1 \atop |i-j|=n} p(i,j) \right\} \quad (7)$$

where: n is the number of grey levels in pixel images, p(i,j) is the normalization GLCM matrix (i,j), and Ng is the maximum number of grey levels in pixel images.

2.2.2.2 Correlation

Correlation is a grey value image, which can determine the contour and displacement.

$$Correlation = \frac{\sum_{i=1}\sum_{j=1}(ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (8)$$

where: μ_x and μ_y are the means of the matrix p_x and p_y; σ_x and σ_y is the standard deviation of the matrix p_x and p_y.

2.2.2.3 Energy

Energy is a measure of the number of repeated pairs.

$$Energy = \sqrt{\sum_i \sum_j \{p(i,j)\}^2} \quad (9)$$

2.2.2.4 Homogeneity

Homogeneity is the similarity of variations in the grey level of the images.

$$Homogeneity = \sum_i \sum_j \frac{1}{1 + (i-j)^2} p(i,j) \quad (10)$$

2.3 Classification Methods

Classification is a process of finding a model to distinguish data classes. This study uses a Support Vector Machine (SVM), Naive Bayes, Multilayer Perceptron (MLP), Multiclass classifier, and Random Forests with Weka machine learning software. Weka provides execution of learning computations that can easily apply to data sets, such as computations for discretization and collation, which can be accessed for free. The workbench combines techniques for fundamental information mining

problems: recurrence, classification, clustering, affiliate rule mining, and quality selection. All calculations take their contributions to the type of solitary social table, which can be read from records or generated by investigation of data sets[6].

2.3.1 Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a directed artificial intelligence calculation that can use for classification and regression. The way SVM works are based on being designed to process data into a hyperplane which classifies the input space into two classes. SVM theory begins with linear grouping cases that can be separated by hyperplane and divided according to their class[26]. Figure 2 shows the principle of SVM on separating two classes using a hyperline.

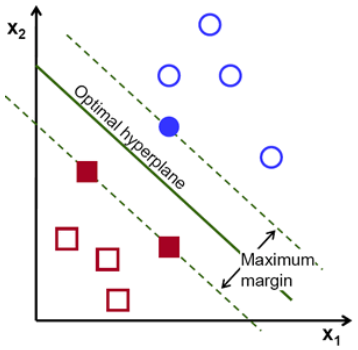


Fig. 2: The principle of SVM on the separation of two classes using a hyperline

2.3.2 Naive Bayes

Naive Bayes is a machine learning model that distinguishes objects based on specific features. Sketchily, Naive Bayes considers that certain features in a class are not related to the presence of other features. This algorithm is advantageous when dealing with large datasets and is easy to build. Bayes's theorem tells how to express terms of quantities that can compute more directly to decide between two classes (e.g. L_1 and L_2), then probabilities (P) for each class is[27]:

$$\frac{P((L_1|feature))}{P((L_2|feature))} = \frac{P(feature|L_1)P(L_1)}{P(feature|L_2)P(L_2)} \quad (11)$$

2.3.3 Multilayer Perceptron (MLP)

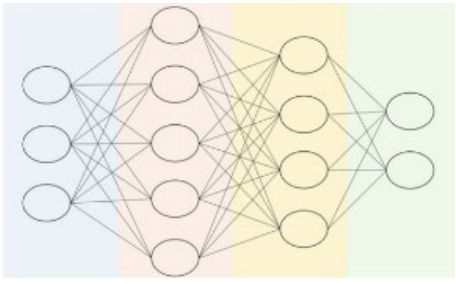


Fig. 3: Basic structure of a multilayer perceptron

Multilayer Perceptron (MLP) is a probabilistic information storage and organization. Perceptron does not follow the given rules (what we call machine programming) but is independently based on the data provided extracted pattern. Neural network is basically in layers organized. Every network has an input and an output layer. Layers between input and output become hidden layers or also called processing layers. Every Layer consists of neurons, which are so-called over-edges joined together[28]. Figure 3 shows the principle of MLP in separating two classes using a network.

2.3.4 Multiclass Classifier

Multiclass classification is a grouping of data when the variable has several classes. It uses a metaclassifier to handle multiclass datasets with a 2-class classifier. This classifier can also apply error correction to improve accuracy further. If the base classifier cannot handle the test load and the test load is not uniform, the information is re-sampled by weight before being passed on to the base classifier. Meta-learning calculations take classifiers and turn them into powerful learning. One parameter indicates the primary classifier; others specify the number of cycles for iterative schemes such as layoffs and favours and the underlying seed for irregular number generators[6].

2.3.5 Random Forest

Random forest is a classification that uses decision trees. Each tree makes predictions independently. The values are averaged using (regression) divided by the selected maximum value (classification) to arrive at the final result. The advantage of this model lies in generating different trees with different subcharacteristics of these traits. Feature selected for each tree is random, so the tree is not too deep and focuses only on feature set. Eventually, when taken together, decision trees provide well-studied predictions[29].

2.4 Data Analysis Technique

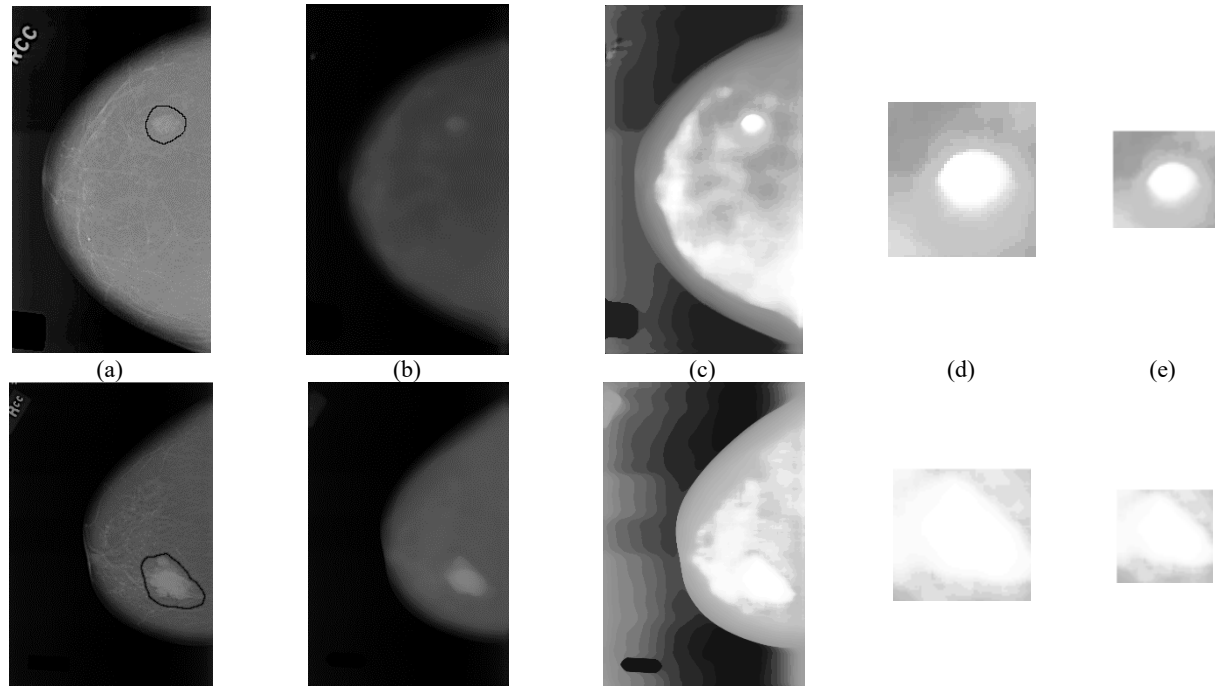
This study performed the data analysis technique using the accuracy, sensitivity, and specificity values of the confusion matrix obtained by the Weka software. The confusion matrix is a benign case that is perceived as benign (True Positive/TP), a benign case that is perceived as malignant (False Positive/FP), a malignant case that is perceived as malignant (True Negative/TN), and a malignant case detected as benign (False Negative/FN). The result is then incorporated into equation[6]:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (13)$$

$$Specificity = \frac{TN}{TN + FP} \times 100\% \quad (14)$$

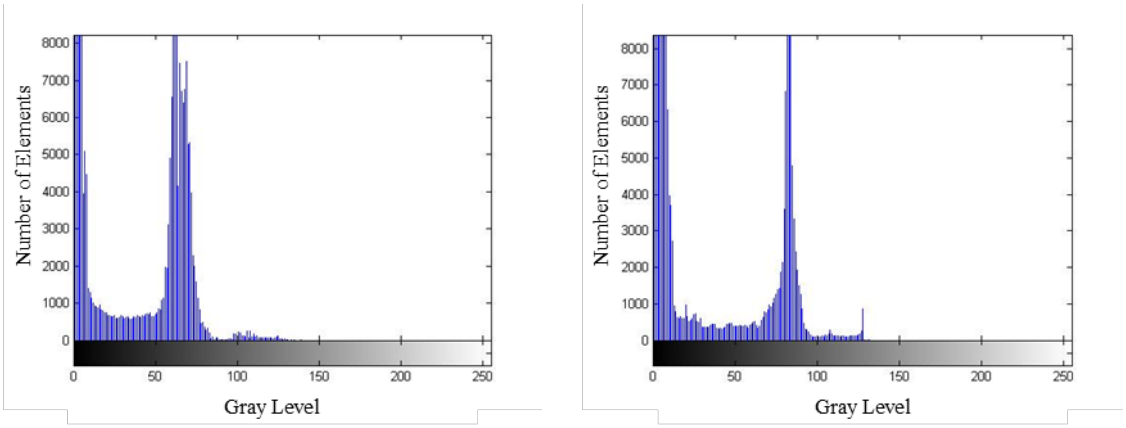
3. Result and Discussion



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Fig 4: Mammography image of the benign (up) and malignant (down) in preprocessing at (a) original; (b) filtering; (c) contrast; (d) cropping; (e) resizing; (f) original; (g) filtering; (h) contrast; (i) cropping; and (j) resizing.



(a) (b)

Fig 5: Histogram image of (a) Benign and (b) Malignant Case

The Results of the preprocessing image in this research are shown in Figure 4. The preprocessing stage has been effectively done, as displayed in the last picture, which will utilize for texture feature extraction. It is effectively eliminated the marker on the tumour mass. Preprocessing also effectively expanded the contrast. The mass was effectively recovered and the picture size was effectively uniformized to 50 x 50 pixels.

The histogram image is shown in Figure 5. The histogram of benign case (Fig 5: (a)) showed high number of elements in gray levels 0-10 and 60-70, meanwhile malignant case (Fig 5: (b)) showed high number of elements in gray level 0-10 and 80-90. The results showed that mammogram images of malignant cases have a higher or lighter gray level than those of benign cases. This difference in gray level is used to calculate first-order statistics.

Mammographic images that have undergone the preprocessing stage are ready for texture feature extraction. The consequences of texture feature extraction are proven in Figure 6. Texture feature data is introduced in ten scatter graphs which aim to see the distribution of the data. The graph shows the average value of each texture feature attribute as a straight line and the average value as a legend. Red is a malignant case, while blue is a benign case.

Figure 6 show tthat extraction results of mean, standard deviation, variance, entropy, kurtosis, contrast, and energy values were higher in the malignant case than in the benign case. Meanwhile, skewness, correlation, and homogeneity values are higher for benign than malignant. This indicates that while the heterogeneity of malignant

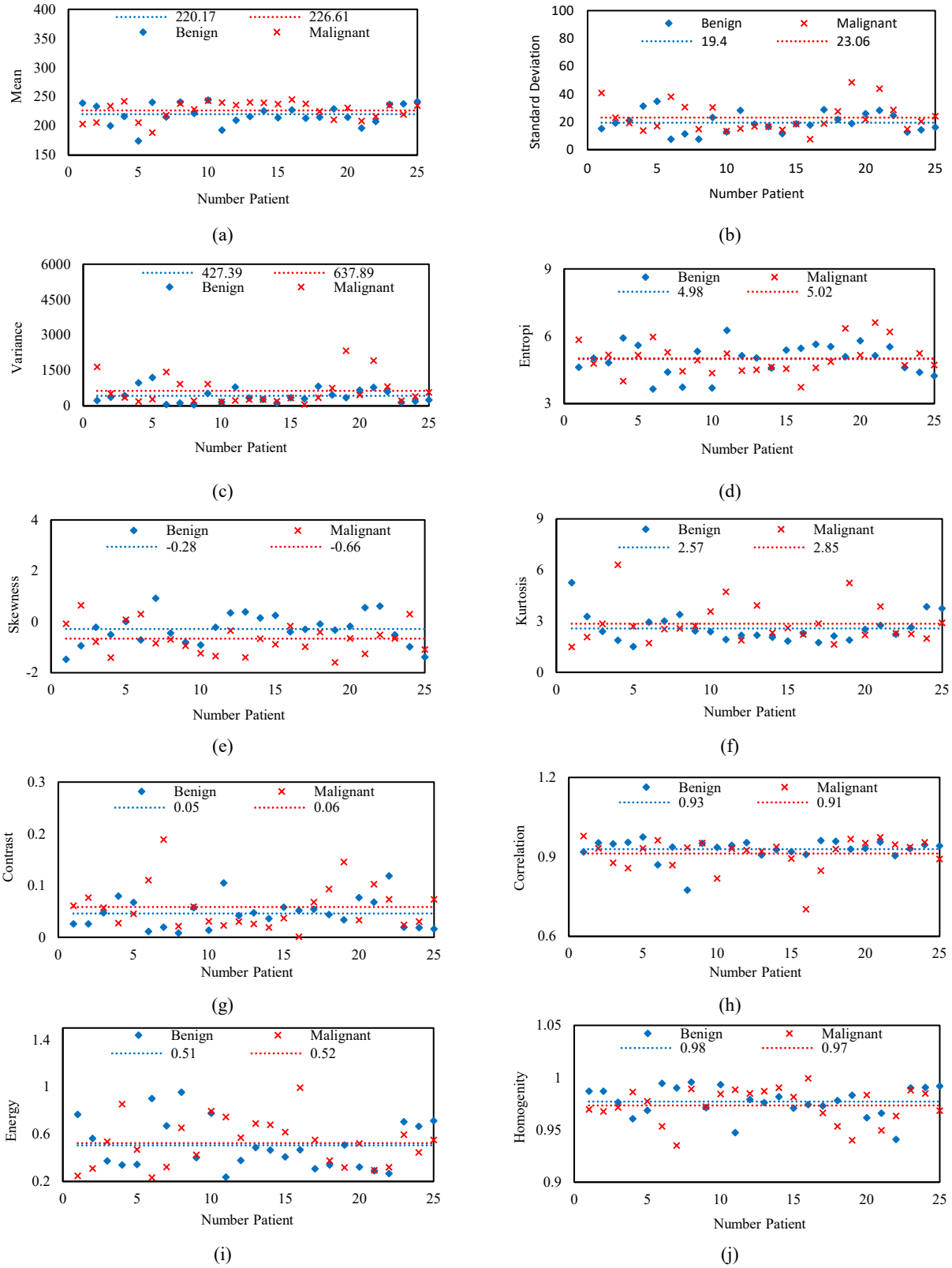


Fig 6: Feature texture extraction at attribute (a) mean, (b) standard deviation, (c) variance, (d) entropy, (e) skewness, (f) kurtosis, (g) contrast, (h) correlation, (i) energy, and (j) homogeneity

cases is higher than that of benign cases, a more uniform and uniform appearance can be observed in benign cases.

Figure 6 show that mean attributes, standard deviation, variance, entropy, kurtosis, contrast, and energy are higher in malignant than benign cases. This indicates that there are deeper grooves in the image texture, and the grey scale

is more complicated in the malignant case. These results are consistent with the relatively complex nature of malignant tumours. In addition, malignant cell size, morphology, and inconsistent polarity characteristics that provide the pathological basis for texture changes often coexist with haemorrhage, discovery, and necrosis [30]. Meanwhile, skewness, correlation, and homogeneity attributes were higher in benign cases than in malignant cases. This shows that the grey level distribution is relatively uniform, the grey level contrast is relatively low, and the local texture regularity is in the benign case.

Subsequently, the texture feature extraction results were used for data classification using Weka software. The confusion matrix and the values of accuracy, sensitivity, and specificity of the five selected classification methods are shown in Table 1. The most accurate data classification was generated by the Random Forest method with 100% accuracy, followed by MLP with 88% accuracy, and the third rank using the Multiclass classifier method with 82%. Likewise the results of sensitivity and specificity. This shows that the random forest method is the best algorithm for distinguishing benign and malignant cases on mammography images.

Based on Table 1, SVM is a data separation technique that finds the optimal hyperplane while maximizing class spacing. To separate classes can use hyperlane features. The position was in the middle between the two classes. The distance between the hyperplane and the data object differs from that of the adjacent (outermost) class. Support vectors are the most difficult to classify because they almost overlap with other classes[31,32]. This causes low accuracy if the condition of the classified data has a more significant gap. In this study, the texture features generated in benign and malignant images have a gap that is not too large, so the SVM method is not appropriate for classifying them. The naive Bayesian because it assumes that the occurrence of a particular function is independent of the occurrence of other functions. One of the main drawbacks of the Naive Bayes classification is its vital independence feature. In practice, having set of entirely independent features is almost impossible[33–36]. In classifying benign and malignant image data on mammography, the texture features are not entirely independent and even overlap. This causes the classification using Naive Bayes to be not optimal.

Based on Table 1, MLP is a neural network that connects many layers of directed graphs. There are only signal paths through the node in one direction. MLP is a deep learning technique because neurons have multiple layers. MLP is supervised machine learning that can solve problems that are not linearly separable, but MLP is sensitive to functional scaling. Feature scaling is a method of having numerical data in a dataset with the same values range. MLP is proven to be able to separate datasets well with up to 100% accuracy[35–37]. The results of texture feature extraction on benign and malignant images in this study show a range that is not too far away so that some similar data are grouped into the same scale. However, this method is quite effective for separating mammographic images in benign and malignant cases, with an accuracy of 88%.

Based on Table 1, unlike the process of binary classification problems, Multiclass classification does not require the selection of scoring thresholds to make predictions. Predicted date is the class with the highest predicted value. The metrics used for Multiclass classification are the same as those used for binary classification. After all the other classes are grouped as belonging to the second class, the metric for each class is calculated by treating it as a binary classification problem[38]. In this study, the texture features on mammographic images of benign and malignant cases have the highest score difference, which is not too significant, resulting in an accuracy rate of 82%.

Based on Table 1, the Random Forest algorithm increases the randomness of the model as the tree grows. Random forest generates multiple decision trees and combines them to get more stable and accurate predictions. A "forest" created from random forests is usually a collection of decision trees trained using the bagging method. Random forests look for the best features in a random subset of features instead of looking for the most critical features when splitting nodes. As a result, this method creates great diversity and generally leads to better models. Using more trees will affect the accuracy. Deciding on a random forest classification based on the voting results of the formed trees will reduce errors[39,40]. The results show an accuracy up to 100% based on texture features to distinguish benign and malignant cases in mammographic images. This is in accordance with previous research which compared several data classification methods which showed that random forest had the highest accuracy[41].

Previously, Kumar et al.[42] conducted a study using MIAS digital mammography data consisting of 322 benign and malignant images. The mammographic image is then cropped using ROI and extracted using GLCM texture features, such as contrast, correlation, energy, and homogeneity. After texture feature extraction, the image is classified using the random forest method, and the results obtained are an accuracy of 97.50%. However, this study showed smaller accuracy results than the current research based on Table 1. It is caused by the more rigid preprocessing in the current study, such as filtering and resizing processes, which can remove noise, artifacts, and homogenize the size of the image. In addition, this study also uses more feature extraction methods, first-order (histogram) and second-order (GLCM), so that more features can be used for the classification process using random forest. The characteristics of the histogram have a significant influence because of the different gray values that can be seen in Figure 5.

Machine learning is a better way of classification compared to conventional machine learning predictive

Table 1. Confusion matrix from Weka software and the result of the accuracy, sensitivity, and specificity.

Classification Methods	Support Vector Machine	Naive Bayes	Multilayer Perceptron	Multiclass Classifier	Random Forest
TP	17 image	19 image	21 image	20 image	25 image
FP	8 image	6 image	4 image	5 image	0 image
TN	14 image	12 image	23 image	21 image	25 image
FN	11 image	13 image	2 image	4 image	0 image
Accuracy	62.00 %	62.00 %	88.00 %	82.00 %	100.00 %
Sensitivity	68.00 %	76.00 %	84.00 %	80.00 %	100.00 %
Specificity	63.64 %	66.67 %	85.19 %	80.77 %	100.00 %

*TP is a benign case that is detected as benign, FP is a benign case that is detected as malignant, TN is a malignant case that is detected as malignant, and FN is a malignant case detected as benign.

models[43] succeeded in predicting the mechanical properties of structural information[44] and overcome the problem of uncertainty[45]. However, this research has limitations in data processing which is still done manually, so it still cannot be implemented practically in health facilities. Nevertheless, the results of a combination of histogram, GLCM, and random forests are capable of achieving an accuracy rate of up to 100%, so it is feasible to be further developed so that it can be applied easily. To achieve this, it can be done by creating an automated system for pre-processing and feature extraction. Using Weka machine learning software, the Random forest can be used as the most optimal classification method to distinguish benign and malignant cases based on texture features in mammography images. This can help radiologists and medical professionals to diagnose cases and take further steps, such as therapy. This can increase the economic and social value of science in the diagnostic process using Weka machine learning which is easy and cheap. However, Weka machine learning cannot display accuracy, sensitivity and specificity values. This software only displays TP, FP, TN, FN and prediction results.

4. Conclusion

A study was conducted to classify benign and malignant cases of mammography images using 25 benign and 25 malignant images. Classification based on texture features with histogram and GLCM (Gray-Level Co-occurrence Matrix) using Weka machine learning software. The result show that the Support Vector Machine (SVM) has an accuracy of 62.00%, sensitivity of 68.88%, and a specificity of 63.64%, Naive Bayes has an accuracy of 62.00%, sensitivity of 76.00%, and a specificity of 66.67%, Multilayer Perceptron (MLP) has an accuracy of 88.00%, sensitivity of 84.00%, and a specificity of 85.19%, Multiclass classifier, has an accuracy of 82.00%, sensitivity of 80.00%, and a specificity of 80.77% and Random Forest method has an accuracy of 100.00%, sensitivity of 100.00%, and a specificity of 100.00% Combination of histogram, GLCM, and random forests are capable of achieving an accuracy rate of up to 100%, so it is feasible to be further developed so that it can be applied easily. This can help radiologists and medical professionals to diagnose breast tumour cases and take further steps, such as therapy. This can increase the economic and social value of science in the diagnostic process using Weka machine learning which is easy and cheap.

References

1) Kemenkes, "Kementerian kesehatan republik indonesia," *Kementeri. Kesehat. RI*, 1 (2019). https://www.kemkes.go.id/article/view/19093000001/penyakit-jantung-penyebab-kematian-terbanyak-ke-2-di-indonesia.html.
2) V. Reddy Allugunti, "Breast cancer detection based on thermographic images using machine learning and deep learning algorithms healthcare view project breast cancer detection based on thermographic images using machine learning and deep learning algorithms," *Comput. Sci.*, **4** (*1*) 49–56 (2022). https://www.researchgate.net/publication/361228083.
3) A.P. Adedigba, S.A. Adeshina, and A.M. Aibinu, "Performance evaluation of deep learning models on mammogram classification using small dataset," *Bioengineering*, **9** (*4*) (2022). doi:10.3390/bioengineering9040161.
4) S. Duffy, D. Vulkan, H. Cuckle, D. Parmar, S. Sheikh, R. Smith, A. Evans, O. Blyuss, L. Johns, I. Ellis, P.E. Sasieni, C. Wale, J. Myles, and S. Moss, "Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the uk age rct," *Health Technol. Assess. (Rockv).*, **24** (*55*) 1–24 (2020). doi:10.3310/hta24550.
5) P.C.B. Sahand Hooshmand, Warren M Reed, Mo'ayyad E Suleiman, "Screening mammography_ diagnostic efficacy—issues and considerations for the 2020s," *Radiat. Prot. Dosimetry*, **197** (*1*) 54–62 (2021). doi:https://doi.org/10.1093/rpd/ncab160.
6) E. Frank, M.A. Hall, and I.H. Witten, "The weka workbench," *Data Min.*, 553–571 (2017). doi:10.1016/b978-0-12-804291-5.00024-6.
7) R. Anand, T. Shanthi, R.S. Sabeenian, M.E. Paramasivam, and K. Manju, "GLCM Feature-Based Texture Image Classification Using Support Vector Machine," in: 3rd EAI Int. Conf. Big Data Innov.

Sustain. Cogn. Comput., 2022: pp. 143–154. doi:10.1007/978-3-030-78750-9_10.
8) M. Euler-chelpin, M. Lillholm, I. Vejborg, M. Nielsen, and E. Lynge, "Sensitivity of screening mammography by density and texture : a cohort study from a population-based screening program in denmark," *Breast Cancer Res.*, **3** 1–7 (2019).
9) S. Selvarajah, and S.R. Kodituwakku, "Analysis and comparison of texture features for content based image retrieval," *Int. J. Latest Trends Comput.*, **2** (*1*) 108–113 (2016).
10) N. Fatima, L. Liu, S. Hong, and H. Ahmed, "Prediction of breast cancer, comparative review of machine learning techniques, and their analysis," *IEEE Access*, **8** 150360–150376 (2020). doi:10.1109/ACCESS.2020.3016715.
11) T. Pang, J.H.D. Wong, W.L. Ng, and C.S. Chan, "Deep learning radiomics in breast cancer with different modalities: overview and future," *Expert Syst. Appl.*, **158** (2020). doi:10.1016/j.eswa.2020.113501.
12) S.J.A.S.F.U.F.A. Bachtiar, "Mammogram Breast Cancer Classification Using Gray-Level Co-Occurrence Matrix and Support Vector Machine," in: 2018 Int. Conf. Sustain. Inf. Eng. Technol., 2018.
13) M. Desai, and M. Shah, "Clinical ehealth an anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (mlp) and convolutional neural network (cnn)," *Clin. EHealth*, **4** (*2021*) 1–11 (2022). doi:10.1016/j.ceh.2020.11.002.
14) Burcu Bektaş; İlkim Ecem Emre; Elif Kartal; Sevinc Gulsecen, "Classification of Mammography Images by Machine Learning Techniques," in: 2018 3rd Int. Conf. Comput. Sci. Eng., IEEE, 2018.
15) J. Dabass, M. Hanmandlu, and R. Vig, "Informatics in medicine unlocked multi-class classification of mammograms with hesitancy based hanman transform classifier on pervasive information set texture features," *Informatics Med. Unlocked*, **26** 100756 (2021). doi:10.1016/j.imu.2021.100756.
16) R. Fadil, A. Jackson, B.A. El Majd, H. El Ghazi, and N. Kaabouch, "Classification of microcalcifications in mammograms using 2d discrete wavelet transform and random forest," *IEEE Int. Conf. Electro Inf. Technol.*, **2020-July** 353–359 (2020). doi:10.1109/EIT48999.2020.9208290.
17) C. Bhowmik, G.S. Pradeep Ghantasala, and R. AnuRadha, "A comparison of various data mining algorithms to distinguish mammogram calcification using computer-aided testing tools," *Lect. Notes Networks Syst.*, **166** 537–546 (2021). doi:10.1007/978-981-15-9689-6_58.
18) A. Arunika, J.F. Fatriansyah, and V.A. Ramadheena, "Detection of asphalt pavement segregation using machine learning linear and quadratic discriminant analyses," *Evergreen*, **9**(*1*) 213–218 (2022). doi:10.5109/4774236.
19) T.G. Patil, "Artificial neural based quality assessment of guava fruit," *Evergreen*, **9**(*2*) 389–395 (2022). doi:10.5109/4794164
20) M. Faid, M. Jasri, and T. Rahmawati, "Perbandingan kinerja tool data mining weka dan rapidminer dalam algoritma klasifikasi," *Teknika*, **8** (*1*) 11–16 (2019). doi:10.34148/teknika.v8i1.95.
21) M. Heath, K. Bowyer, D. Kopans, R. Moore, and R. Kegelmeyer Jr., "USF digital mammography home page," *Proc. Fifth Int. Work. Digit. Mammogr.*, 212–218 (2001).
22) I.A. Brizuela Pineda, R.D. Medina Caballero, J.J. Cáceres Silva, J.C. Mello Román, and J.L. Vázquez Noguera, "Quadri-histogram equalization using cutoff limits based on the size of each histogram with preservation of average brightness," *Signal, Image Video Process.*, **13** (*5*) 843–851 (2019). doi:10.1007/s11760-019-01420-9.
23) R. Javed, M.S.M. Rahim, T. Saba, S.M. Fati, A. Rehman, and U. Tariq, "Statistical histogram decision based contrast categorization of skin lesion datasets dermoscopic images," *Comput. Mater. Contin.*, **67** (*2*) 2337–2352 (2021). doi:10.32604/cmc.2021.014677.
24) L. Losurdo, A. Fanizzi, T.M.A. Basile, R. Bellotti, U. Bottigli, R. Dentamaro, V. Didonna, V. Lorusso, R. Massafra, P. Tamborra, A. Tagliafico, S. Tangaro, and D. La Forgia, "Radiomics analysis on contrast-enhanced spectral mammography images for breast cancer diagnosis: a pilot study," *Entropy*, **21** (*11*) (2019). doi:10.3390/e21111110.
25) M. Hall-Beyer, "GLCM texture: a tutorial v. 3.0," *Arts Res. Publ.*, (*2017–03*) 75 (2017). https://prism.ucalgary.ca/handle/1880/51900%0Ahttp://hdl.handle.net/1880/51900.
26) A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, "Sentimen analisis publik terhadap kebijakan lockdown pemerintah jakarta menggunakan algoritma svm," *Jdmsi*, **2** (*1*) 31–37 (2021). https://t.co/NfhnfMjtXw.
27) Y. Zamrodah, "Bayesian Artificial Intelligence (Second Edition)," CRC Press, 2016.
28) D. Sonnet, V. Perceptron, and D. Learning, "Neuronale Netze kompakt," Springer Fachmedien Wiesbaden GmbH, Hamburg, 2022.
29) W. Sullivan, "Machine learning for beginners algorithms, decision tree random forest introduction," (2017).
30) Q. Zhao, C.Z. Shi, and L.P. Luo, "Role of the texture features of images in the diagnosis of solitary pulmonary nodules in different sizes," *Chinese J. Cancer Res.*, **26** (*4*) 451–458 (2014). doi:10.3978/j.issn.1000-9604.2014.08.07.
31) T. Sihotang, and M. Rafli, "Classification of book types using the support vector machine (svm) method," *J. Mantik*, **6** (*36*) 43–49 (2022).
32) S. Raghavendra, and P.C. Deka, "Support vector

machine applications in the field of hydrology: a review," *Appl. Soft Comput. J.*, **19** 372–386 (2014). doi:10.1016/j.asoc.2014.02.002.
33) M.B.& S.P. Yang Can Gong, "Texture Based Mammogram Classification and Segmentation," in: IWDM 2006, Digital Mammography, 2006: pp. 616–625.
34) E. Ramadhani, I. Husein, I. Zufria, H. Nasution, U.S. Kuala, P.S. Informatika, U. Islam, N. Sumatera, U. Islam, N. Sumatera, and U.N. Medan, "Naive bayes algorithm to predict student success in a course," *Int. J. Adv. Sci. Technol.*, **29** (*8*) 2964–2974 (2020).
35) K. Knezevic, J. Fulir, D. Jakobovic, and S. Picek, "NeuroSCA : evolving activation functions for side-channel analysis," *Int. Assoc. Cryptologic Res.*, 1–23 (2021).
36) G.E. Pintelas, "Deep neural networks for cryptocurrencies price prediction," *Deep Neural Networks Bitcoin Price Predict.*, (*June*) 73 (2020). https://d-nb.info/1185667245/34.
37) L. Istianah, and H. Sumarti, "Classification of pneumonia in thoracic x-ray images based on texture characteristics using the mlp (multi-layer perceptron) method," *J. Nat. Sci. Math. Res.*, **6** (*2*) 78–84 (2020).
38) M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," 1–17 (2020). http://arxiv.org/abs/2008.05756.
39) M.A. Berawi, S.A.O. Siahaan, Gunawan, P. Miraj, and P. Leviakangas, "Determining the prioritized victim of earthquake disaster using fuzzy logic and decision tree approach," *Evergreen*, **7**(*2*) 246–252 (2020). doi:10.5109/4055227.
40) A. Alkazaz, and M.S. Kharouki, "Forest algorithm for classification evaluation of adaptive random forest algorithm for classification of evolving data stream," (2020).
41) M. Pradesh, and C. Engineering, "Design and development of machine learning assisted cylindrical dielectric resonator antenna," *Evergreen*, **10**(*1*) 308–316 (2023). doi:10.5109/6781085
42) T. Ananth Kumar, G. Rajakumar, and T.S. Arun Samuel, "Analysis of breast cancer using grey level co-occurrence matrix and random forest classifier," *Int. J. Biomed. Eng. Technol.*, **37** (*2*) 176–184 (2021). doi:10.1504/IJBET.2021.119503.
43) K. Sujatha, N.P.G. Bhavani, V.S. George, T.K. Reddy, N. Kanya, and A. Ganesan, "Innovation in agriculture industry by automated sorting of rice grains," *Evergreen*, **10**(*1*) 283–288 (2023). doi:10.5109/6781076.
44) J.F. Fatriansyah, S.N. Surip, and F. Hartoyo, "Mechanical property prediction of poly(lactic acid) blends using deep neural network," *Evergreen*, **9**(*1*) 141–144 (2022). doi:10.5109/4774229.
45) G.D. Nugraha, B. Sudiarto, and K. Ramli, "Machine learning-based energy management system for prosumer," *Evergreen*, **7**(*2*) 309–313 (2020). doi:10.5109/4055238.