Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics e-ISSN: 2656-8624
Homepage: https://ijeeemi.org/; Vol. 7, No. 4, pp. 673-686, November 2025

RESEARCH PAPER **OPEN ACCES**

Hyperparameter optimization of breast ultrasound image classification models using ant colony optimization based on texture features

Ahmad Fauzi[1], Lukmanda Evan Lubis[2], Raju Wandira[1], and Syarto Musthofa[3]

[1] Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia
[2] Departement of Physics, Universitas Indonesia, Depok, Indonesia
[3] Mathematics Study Program, UIN Imam Bonjol Padang, Padang, Indonesia

ABSTRACT

Breast cancer is one of the most prevalent types of cancer in humans and carries the highest cumulative risk compared to other cancers. Accurate diagnosis and efficient intervention for this disease are very important to improving patient survival. However, the accuracy of these diagnostic assessments often depends on the subjectivity and expertise of pathologists, which may lead to inconsistencies. To address this issue, automated diagnosis using machine learning has been developed. This study aims to optimize machine learning algorithms using texture features to produce an efficient breast cancer classification model that remains competitive with deep learning-based models. The main contribution of this research lies in enhancing the performance of machine learning classifiers for breast ultrasound image classification and providing insights into the effectiveness of optimization across different algorithms. The dataset used in this study consists of three classes (benign, malignant, and normal) with a total of 780 breast ultrasound images from 600 patients. All images were processed and augmented to increase data variation before modeling with five classification algorithms: Random Forest, SVM, Decision Tree, Gradient Boosting, and k-NN. Modeling was conducted in two scenarios: one without optimization and the other with hyperparameter optimization using the Ant Colony Optimization algorithm. The results showed that the GLCM angle orientation had a relatively small effect on model performance. The best accuracy for each orientation was achieved with k-NN+ ACO (0.95 at 0^0), SVM+ACO (0.94 at 45^0), SVM+ACO (0.90 at 90^0), and RF+ACO (0.95 at 135^0). In conclusion, classification using five GLCM-extracted features optimized with ACO achieved better performance than approaches without optimization or those previously applied using deep learning methods in other studies.

Paper History
Received August 02, 2025
Revised October 20, 2025
Accepted October 25, 2025
Published October 28, 2025

Keywords
Breast cancer;
Machine learning;
Texture features;
Ant Colony Optimization;
Deep learning

Author Email
ahmadfauzi@uinib.ac.id
lukmanda.evan@sci.ui.ac.id
rajuwandira@uinib.ac.id
syartom@uinib.ac.id

I. Introduction

Breast cancer is one of the most common cancers among women worldwide. According to data from the World Health Organization (WHO), in 2022, the number of breast cancer cases among women increased by more than 2 million, making it the second most common cancer affecting humans and the one with the highest cumulative risk compared to other types of cancer [1]. Despite its high prevalence, accurate diagnosis and effective intervention for this disease can significantly improve patient survival [2]. In general, diagnoses made by pathologists to determine the stage of a patient's cancer are based on medical imaging results, such as mammography, biopsy (histopathology), or ultrasound. However, the accuracy of these diagnoses is highly dependent on the subjectivity and expertise of the pathologist, which may lead to inconsistencies. To address this issue, diagnostic assistance systems, such as Computer-Assisted Diagnosis (CAD), have been developed to enhance both efficiency and accuracy in disease diagnosis, including cancer. The current development of CAD has been influenced by the field of artificial intelligence, specifically machine learning, which constitutes one of its major branches. Using the Support Vector Machine (SVM), Aprilia et al. [3] classified breast cancer ultrasound images into two classes, benign and malignant.

The features used in their study were extracted using the Gray Level Co-Occurrence Matrix (GLCM). Their research demonstrated that the SVM was capable of classifying breast cancer ultrasound images with a maximum accuracy of 79%, both in unbalanced and balanced classes. In addition, using the Wisconsin Breast Cancer Diagnostic dataset from the UCI Machine Learning Repository, Minnoor et al. [4] developed a breast cancer classification model based on Random Forest (RF). Using a similar dataset, Uddin et al. [5] used Principal Component Analysis (PCA) to reduce the feature dimensions from 30 to 16 features. The PCA-derived features were then utilized to construct a model using a machine learning algorithm. In contrast, Aamir et al. [6] reduced features by examining strong correlations between predictors, which were then used to build a breast cancer classification model. Meanwhile, a study by Kaunang et al. [7] comparing RF with Decision Tree (DT)

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

for breast cancer detection showed that RF achieved 9.5% higher accuracy than DT.

In addition to machine learning, deep learning has been widely adopted in CAD development, particularly for medical image classification. Unlike machine learning, which depends on handcrafted features, deep learning automatically extracts features through a convolution process. Using data from Baheya Hospital [8], Pacal [9] evaluated the performance of several well-known CNN architectures (AlexNet, ResNet, VGG, GoogleNet, and EfficientNet) for breast cancer classification. Meanwhile, Surya et al. [10] utilized the MobileNetV2 architecture to enhance efficiency while maintaining classification accuracy. A different approach was proposed by Cruz-Ramos et al. [11], who combined features extracted from a CNN with handcrafted features, achieving better performance compared to CNN models that rely solely on automatic features. To further increase accuracy in breast cancer histopathology images, Istighosah et al. [12] implemented the ResNet101 architecture, while another study [13] applied Particle Swarm Optimization (PSO) to optimize the weights of the final dense layer in the CNN model.

Although deep learning has shown better performance than machine learning [14][15][16], this approach requires high computational resources [17] and long computation time [18]. Therefore, this study aims to optimize machine learning algorithms using a limited number of features derived from texture features extracted using the GLCM, resulting in an efficient breast cancer classification model that remains competitive with models developed through deep learning.

The GLCM has been widely applied in various medical imaging, such as diabetic retinopathy [19][20][21], lung cancer [22][23], cervical cancer [24], colorectal polyps [25], tendon imaging [26], breast cancer [3][27][28], and brain tumor analysis [29]. This approach captures spatial relationships between neighboring pixels in an image. Through this approach, both micro and macro-level texture patterns can be effectively identified.

The main contributions of this study are as follows: First, it enhances the performance of machine learning classifiers for breast cancer ultrasound image classification. Second, it develops a texture-based classification model that demonstrates competitive performance compared to deep learning approaches. Third, it provides a comparative analysis of optimization times across several machine learning algorithms. Ultimately, it provides valuable insights into the effectiveness of optimization algorithms in enhancing classifier performance for breast cancer detection.

This study is structured as follows. Section 2 presents the proposed method, the dataset used, the preprocessing steps, the feature extraction process, the handling of imbalanced data, and the models along with their evaluation metrics. Section 3 reports the experimental results, including accuracy, precision, recall, F1-score, and the computational time required for optimization. Section 4 provides an interpretation and comparison of the results with related studies and discusses the study's limitations. Finally, Section 5 concludes the paper by highlighting the main findings and outlining directions for future research.

II. Materials and Methods

In this study, we used Python 3.13 with the NumPy and scikit-learn libraries. NumPy was employed to construct and manipulate multidimensional arrays representing the images, while scikit-learn was used for data classification. The experiments were conducted on a macOS system with an Intel Core i5 processor at 2.5 GHz and 15 GB of DDR3 RAM.

A. Dataset

The method proposed in this study can be seen at Fig 1. In this study, we utilized the Breast Ultrasound Images (BUSI) dataset [8], which comprises breast ultrasound images from 600 female patients aged 25 to 75 years. Al-Dhabyani and Gomaa originally collected the dataset at Baheya Hospital (Egypt) and contains 437 benign images, 210 malignant images, and 133 normal images [8]. The dataset consists of three classes (normal,

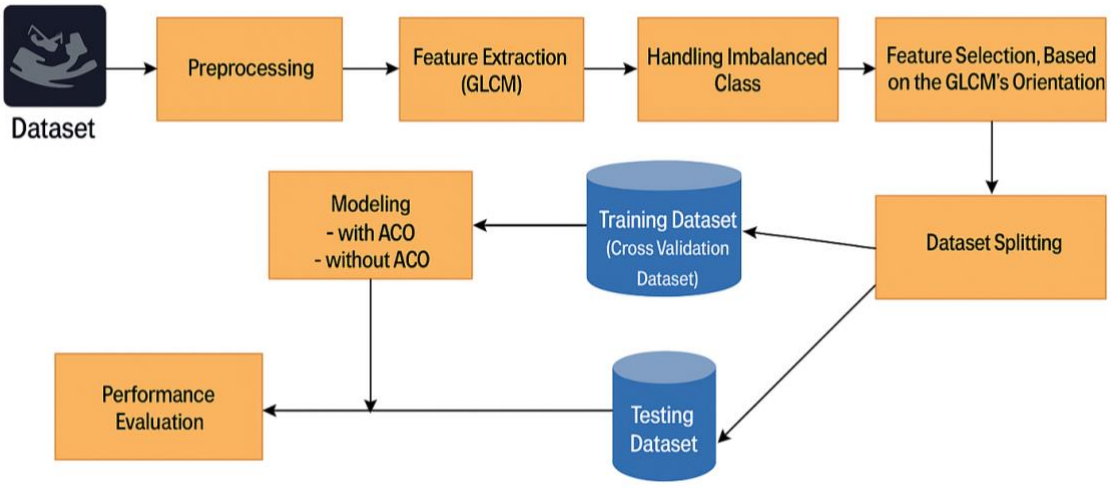

Fig 1. Proposed method

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

malignant, and benign) with an average size of 500x500 pixels in PNG format. Data acquisition was performed using the LOGIQ E9 ultrasound system and the LOGIQ E9 agile ultrasound system.

B. Data Processing

The data pre-processing stage in this study began by standardizing the size of all images from various classes to 224x224 pixels, ensuring consistent input to the machine learning model. Since the GLCM features require single-channel images, the original images were separated into three channels (red, green, and blue). Among these, only the green channel was used for feature extraction, as several previous studies have demonstrated that it provides better contrast and lower noise levels than red and blue channels [30], making it more suitable for medical image texture analysis. In addition to resizing and channel separation, data augmentation was applied to the green channel to enrich image variation and reduce the risk of overfitting during the modelling process [31]. Ramamoorthy et al. [32] also reported that data augmentation can enhance the performance of classification models. The augmentation techniques used in this study included horizontal and vertical flipping, and rotations of $+90^0$ and -90^0 (Fig 2). As a result, the number of images at the modeling stage increased to 2185 for the benign class, 1050 for the malignant class, and 665 for the normal class.

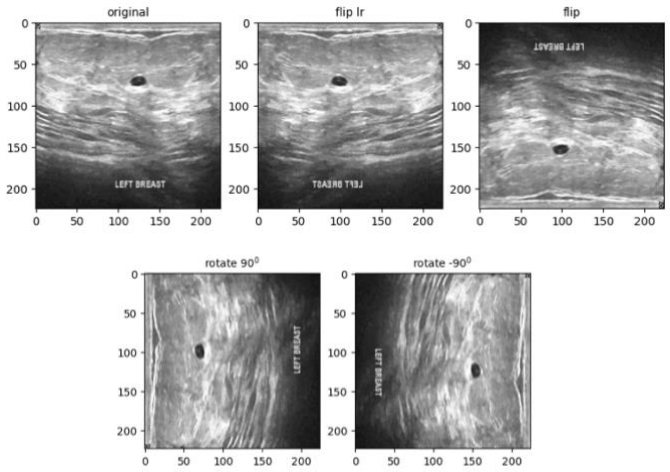

Fig 2. Data augmentation: Horizontal flipping, vertical flipping, 90^0 rotation, and -90^0 rotation

C. Feature Extraction

This study utilized texture features derived from the GLCM, which captures texture regularity by quantifying the frequency of pixel intensity pairs at defined distances and orientations [33]. As the images are grayscale, second-order statistical analysis is more efficient since it emphasizes luminance (intensity) rather than color. In this research, the GLCM was computed by considering the distance between pixels of d=1, and four angular orientations (0^0, 45^0, 90^0, and 135^0). These orientations were selected to capture the variation in texture direction that may occur in the images, thereby ensuring a more comprehensive dataset. Unlike the study by Andra [34], which combined all GLCM features, regardless of orientation, into a single model, this study performs classification using GLCM features for each orientation, thereby producing distinct models for each orientation. The features extracted using the GLCM in this work include:

1. Contrast

Contrast (Eq. (1) [33]) is used to measure the intensity difference between a pixel and its neighbors. The higher the contrast value, the greater the local texture difference.

$$\text{Contrast} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-j)^2.P(i,j) \quad (1)$$

where $P(i,j)$ is the probability value or relative frequency of the appearance of pixel pairs with intensities i and j, at a certain distance and angle.

2. Dissimilarity

Dissimilarity (Eq. (2) [35]) is used to measure the degree of dissimilarity between pixel pairs linearly.

$$\text{Dissimilarity} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|i-j|.P(i,j) \quad (2)$$

3. Homogeneity

Homogeneity (Eq. (3) [35]) is used to describe texture uniformity. A high value indicates a more homogenous texture.

$$\text{Homogeneity} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{P(i,j)}{1+(i-j)^2} \quad (3)$$

4. Correlation

Correlation (Eq. (4) [33]) is used to assess the linear relationship between pixel values and their neighbors. The value ranges from -1 to 1.

$$\text{Corr} = \frac{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-\mu_i)(j-\mu_j).P(i,j)}{\sigma_i\sigma_j} \quad (4)$$

where μ_i, μ_j are the mean, and σ_i, σ_j are the standard deviations of the rows and columns of the GLCM.

5. Angular Second Moment (ASM)

ASM (Eq. (5) [33]) is used to indicate the level of smoothness and uniformity of texture. A high value indicates a more homogenous texture.

$$\text{ASM} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}P(i,j)^2 \quad (5)$$

D. Imbalanced Classes

Based on the dataset obtained by Al-Dhabyani and Gomaa [8], the distribution of data among classes was found to be imbalanced. This imbalance can affect the performance of the classification model. The model tends to be biased towards the majority class, thereby reducing its performance on the minority class [36]. To overcome this problem, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to the GLCM-generated features of each minority class, producing new variations that are more representative while reducing the risk of

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

overfitting [25][36]. By applying the SMOTE, the number of samples in each class was balanced, allowing the model to learn patterns from all classes more fairly.

After the balancing process, the dataset was divided into two subsets: 80% for training and 20% for testing. To evaluate the model's performance more accurately and prevent overfitting, 5-fold cross-validation was applied to the training data. In this process, the training set is divided into five subsets. In each iteration, four subsets are used for training, and one is reserved for validation. This procedure is repeated five times so that each subset is used once as validation data, and the results are averaged to obtain the final model performance. Furthermore, during the training and testing stages, the random.The seed function is employed to control the randomization process, ensuring that the data used for model construction remains consistent. Consequently, the experimental results are reproducible and reliable.

E. Modeling and Evaluation

In this study, five classification algorithms were implemented: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting (GB), and k-Nearest Neighbour (k-NN). These algorithms have been widely applied not only for breast cancer [7][36], but also for other medical classification tasks [37][38][39][40].

DT is a supervised learning algorithm that classifies data by recursively splitting it into subsets based on the values of specific features. At each node, the algorithm selects the feature that provides the highest Information Gain (IG), as defined in Eq. (6), which is computed from the Entropy (E) in Eq. (7).

$$IG(S,A) = E(S) - \sum \frac{|S_v|}{|S|} E(S_v) \quad (6)$$

$$E(S) = -\sum_{i=1}^{N} p_i \log_2(p_i) \quad (7)$$

where $IG(S,A)$ is information gain obtained by splitting on feature A, $E(S)$ is entropy of the dataset, $E(S_v)$ is entropy of subset, S_v is number of samples in subset, S is number of samples in the dataset, N is number of classes, and p_i is proportion of samples in class i.

RF is an ensemble learning algorithm that combines multiple DT to improve classification performance and reduce overfitting. Each Decision Tree is trained on a different subset of the dataset, a process known as bootstrap sampling, which helps to minimize the correlation among the trees. The final prediction (y) is determined through a majority voting process based on the classification results of all individual trees (Eq. (8)).

$$y = mod(y_1, y_2, \dots, y_n) \quad (8)$$

SVM is a supervised learning algorithm that uses a hyperplane to separate data into different classes. In addition, SVM employs kernel functions to enable the classification of non-linear data by transforming the input features into a higher-dimensional space where the classes become linearly separable (Eq. (9)).

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \quad (9)$$

where $f(x)$ is the predicted class label of the input x, α_i is the Lagrange multiplier associated with the ith training sample, y_i indicates the actual class label of the ith sample, $K(x_i, x)$ is the kernel function, and b is the bias.

GB is an ensemble learning method that builds a predictive model by sequentially combining multiple weak learners (Eq. (10)). Unlike RF, which builds trees independently and aggregates their outputs by averaging or voting, GB constructs trees sequentially, where each new tree attempts to correct the errors made by the previous ensemble.

$$F(x) = F_0(x) + \sum_{n=1}^{N} \eta h_n(x) \quad (10)$$

where $F(x)$ is the final model, $F_0(x)$ is the initial model, h_n represents the nth tree, η is learning rate ($0 < \eta \leq 1$), and N denotes the total number of trees. k-NN is a non-parametric algorithm in which the class of a data point is determined based on the majority class among its nearest neighbors. Therefore, the distance between the data point and its neighbors, as defined in Eq. (11), plays a crucial role in determining the classification result

$$d(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{\frac{1}{p}} \quad (11)$$

where x and y represent two data points, n is the number of features, and p defines the type of distance metric. If p is 1, the distance is calculated using the Manhattan distance, whereas if p is 2, it is calculated using the Euclidean distance. Furthermore, to evaluate their performance, each algorithm was tested using two scenarios, without parameter optimization (representing the baseline model), and with parameter optimization using ACO. ACO is a metaheuristic algorithm inspired by the behaviour of ant colonies in finding the shortest path to a food source [41]. At each iteration, a number of agents explore the solution space. The selection of a particular solution is based on pheromone intensity (τ), which reflects the collective knowledge of previous solutions, and heuristic information (η), which provides an initial preference for promising solutions. In this study, the number of ants (n_ants=10), the number of iterations (T=10), and the evaporation rate (ρ=0.3) were used.

In contrast to brute-force optimization methods, such as GridSearch, and probabilistic model-based approaches, such as Bayesian optimization, metaheuristic approaches can overcome the problem of local optima and thus identify more optimal solutions [42]. Akbari et al. [43] demonstrated that hyperparameter tuning for identifying prospective regions of porphyry-Cu mineralization using ACO resulted in higher model accuracy than tuning with GridSearch. Similarly, Said et al. [44] reported that metaheuristic approaches outperformed both Bayesian and search-based methods.

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

Moreover, although not always applied for hyperparameter optimization, several studies have demonstrated that ACO can improve model performance more effectively than other metaheuristic approaches. Gjecka and Bushati [45] reported that using ACO for feature selection in thyroid disorder prediction achieved an accuracy of 99.5%, which was significantly higher than that obtained with PSO (93%) and GA (92%). Similarly, Alghawli and Taloba [46] showed that feature selection using ACO for depressive disorder classification produced higher accuracy compared with PSO and GA feature selection. In another comparative study, Adrian et al. [47] found that among the three metaheuristic methods (GA, PSO, and ACO), ACO achieved the most efficient execution time when solving construction site layout problems.

Pseudocode

```
Input: Models M, Hyperparams H, Data D, ACO params (n_ants=10, T=10, α, β, ρ=0.3, Q)
Initialize pheromone τ uniformly (τ0)
for t = 1 to T:
   solutions = []
   for k = 1 to n_ants:
      Select model m_j with prob P_j ∝ τ_j^α * η_j^β
      For each hyperparameter dim in H_j:
         Select value with prob P_h ∝ τ_h^α * η_h^β
      Train model m_j with chosen hyperparams on the train set
      Evaluate on validation set → score_k
      Store solution path and cost L_k = 1/score_k
   Evaporate pheromone: τ ← (1-ρ)*τ
   For each solution:
      Update pheromone: τ_path += Q/L_k
   Update global best if score_k > best_score
Return the best model and hyperparameters
```

This study employed ACO to determine the optimal combination of hyperparameters within predefined ranges. For the RF classifier, the number of estimators (i.e., the number of decision trees in the forest) was set to 50, 100, and 150; the maximum tree depth was tested at None (unlimited), 5, 10, and 20; and the minimum number of samples required to split an internal node was set to 2, 5, and 10. For the DT, the maximum depth was varied across four levels: None, 5, 10, and 20, while the minimum samples per split were set at 2, 5, and 10. For the GB, the number of estimators was set to 50, 100, and 150; the learning rate was tested at 0.01, 0.1, and 0.2; and the maximum depth was set at 3, 5, and 7. For the SVM, the regularization parameter C was set to 0.1, 1, 10, 100, and 1000; the kernel coefficient γ (gamma) was tested at 1e−4, 1e−3, 0.01, 0.1, 1, and 10, with the radial basis function (RBF) kernel. Finally, for the k-NN, the number of neighbors, k, was tested at 2, 5, 8, and 10, with distance metrics defined by the Minkowski parameter p set to 1 (Manhattan distance) and 2 (Euclidean distance).

The final phase of this study involved evaluating the model's performance using test data, which comprised 20% of the total dataset. This evaluation aimed to assess the model's ability to make predictions on unseen data, thereby objectively reflecting the model's performance. Using a confusion matrix, the analysis was conducted by comparing the baseline model (without ACO optimization) with the model optimized using ACO. The confusion matrix, which contains True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), then becomes the basis for calculating evaluation metrics, including accuracy (Eq. (12)), precision (Eq. (13)), recall (Eq. (14)), and F1-score (Eq. (15)). The calculation of these metrics provides a comprehensive overview of the model performance and the impact of ACO on improving predictive performance.

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

$$\text{Precision} = \frac{TP}{TP + FP} \quad (13)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (14)$$

$$\text{F1} - \text{score} = \frac{2.\text{Precision}.\text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

III. Results

A total of 500 randomly selected samples per class were used to assess feature significance through ANOVA and the Kruskal-Wallis test. Consistent across all orientations, the results demonstrated that four features (contrast, dissimilarity, homogeneity, and correlation) significantly contributed to classification (Tables 1–4). In comparison, ANOVA revealed that ASM was not significantly different for class discrimination ($p > 0.05$), whereas the Kruskal-Wallis test identified this feature as highly significant ($p < 0.05$) across all orientations. Therefore, all GLCM-extracted features were considered appropriate for modeling.

Table 1. Significance test of features at 0^0

	ANOVA (p)	Kruskal-Wallis (p)
Contrast	5.03E-07	3.41E-08
Dissimilarity	2.06E-07	4.48E-08
Homogneity	3.92E-07	1.27E-08
Correlation	7.69E-07	2.29E-06
ASM	2.04E-01	6.44E-10

Table 2. Significance test of features at 45^0

	ANOVA (p)	Kruskal-Wallis (p)
Contrast	1.79E-40	9.43E-42
Dissimilarity	1.18E-44	1.60E-41
Homogneity	6.76E-09	9.62E-20
Correlation	9.62E-26	1.86E-28
ASM	1.96E-01	2.20E-12

Table 3. Significance test of features at 90^0

	ANOVA (p)	Kruskal-Wallis (p)
Contrast	3.00E-13	7.51E-12

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

Dissimilarity	2.51E-11	7.10E-11
Homogneity	9.66E-07	3.21E-09
Correlation	4.88E-13	1.02E-10
ASM	2.04E-01	2.28E-10

Table 4 Significance test of features at 135^0

	ANOVA (p)	Kruskal-Wallis (p)
Contrast	4.79E-40	3.45E-41
Dissimilarity	1.06E-43	4.09E-41
Homogneity	7.86E-09	7.10E-20
Correlation	7.07E-25	2.73E-27
ASM	1.96E-01	2.06E-12

The results of the model performance evaluation are presented in Tables 5–8, which report the accuracy of validation results obtained using 5-fold cross-validation. At the 0^0 orientation (Table 5), hyperparameter optimization with ACO significantly improved model performance, particularly for SVM, GB, and k-NN, which increased by 38.43%, 19.88%, and 9.68%, respectively, compared with models without optimization. A similar trend was observed at the 45^0 orientation (Table 6), with accuracy improvements of 34.07% for SVM, 16.91% for GB, and 6.73% for k-NN. At the 90^0 orientation (Table 7), SVM accuracy increased by 38.40%, GB by 19.31%, and k-NN by 8.35%. Likewise, at the 135^0 orientation (Table 8), the accuracy of SVM, GB, and k-NN increased by 35.37%, 18.95%, and 7.23%, respectively. In contrast, RF and DT showed relatively stable performance under both optimized and non-optimized conditions.

The results in Tables 5–8, also demonstrate that changes in the GLCM angle orientation had no significant effect on model performance. RF, SVM, DT, GB, and k-NN consistently maintained the same relative ranking across orientations. These findings suggest that, within the context of ultrasound image datasets, the composition of GLCM features has a greater impact on model performance than the orientation direction in the Co-occurrence matrix. Thus, although angle orientation is part of the GLCM formulation, its role in influencing the performance of classification models in this study was relatively small. Moreover, appropriate hyperparameter optimization has been shown to have a significant impact on model performance.

Table 5. Comparison of the model performance at the 0^0 GLCM orientation

Model Validation	Average of Accuracy Default with Hyperparameter	Average of Accuracy with Tuned Hyperparameter by ACO
RF	0.9159±0.001	0.9184±0.001
SVM	0.5282±0.000	0.9125±0.000
DT	0.8603±0.001	0.8636±0.001
GB	0.7019±0.000	0.9007±0.003
k-NN	0.8105±0.000	0.9073±0.001

Table 6. Comparison of the model performance at the 45^0 GLCM orientation

Model Validation	Accuracy Default with Hyperparameter	Accuracy with Tuned Hyperparameter by ACO
RF	0.9212±0.000	0.9199±0.002
SVM	0.5727±0.000	0.9134±0.002
DT	0.8621±0.001	0.8629±0.001
GB	0.7347±0.000	0.9038±0.004
k-NN	0.8383±0.000	0.9056±0.000

Table 7. Comparison of the model performance at the 90^0 GLCM orientation

Model Validation	Accuracy Default with Hyperparameter	Accuracy with Tuned Hyperparameter by ACO
RF	0.8774±0.002	0.8819±0.003
SVM	0.5046±0.000	0.8886±0.000
DT	0.8257±0.002	0.8303±0.001
GB	0.6693±0.000	0.8624±0.005
k-NN	0.7849±0.000	0.8684±0.000

Table 8. Comparison of the model performance at the 135^0 GLCM orientation

Model Validation	Accuracy Default with Hyperparameter	Accuracy with Tuned Hyperparameter by ACO
RF	0.9240±0.001	0.9275±0.001
SVM	0.5696±0.000	0.9233±0.000
DT	0.8723±0.001	0.8733±0.001
GB	0.7229±0.000	0.9124±0.001
k-NN	0.8354±0.000	0.9077±0.000

Fig. 3 illustrates clear variations in optimization times across the classifiers for 10 iterations, reflecting the differences in computational complexity inherent to each algorithm. The DT achieved the fastest optimization time (0.36 minutes) because it constructs only a single tree with straightforward splitting rules, resulting in minimal computational cost. Similarly, the k-NN required relatively little time (0.88 minutes), as its optimization primarily involves recalculating distances for different parameter values without the need to build an explicit model. In contrast, the RF required more time (12 minutes) since it constructs multiple decision trees within an ensemble, thereby increasing computational demand compared to a single DT. The SVM with an RBF kernel took substantially longer (46 minutes) due to the quadratic optimization

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

process involved in determining the optimal hyperplane, which becomes more computationally intensive when kernel transformations are applied. The GB recorded the longest optimization time (95 minutes) because it sequentially builds trees, with each tree correcting the errors of the previous one, making the process both computationally expensive and less parallelizable.

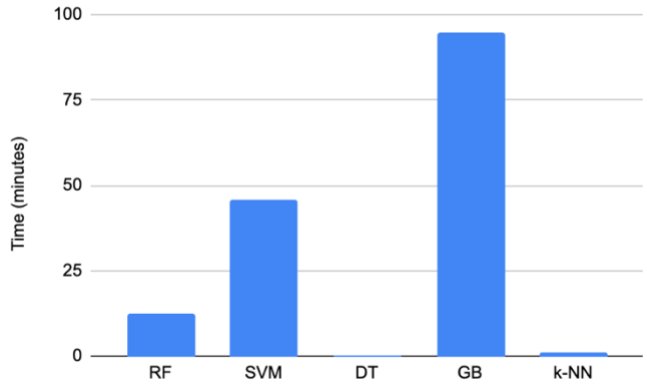

Fig 3. Comparison of optimization time over 10 iterations for each classifier

Similar to the validation results, the evaluation on the testing data demonstrated that the ACO-optimized models achieved a significant improvement in performance compared with the non-optimized models (Table 9). At the 0° orientation, the best performance was achieved by k-NN with ACO, reaching an accuracy of 95%, with precision, recall, and F1-score of 0.96, 0.96, and 0.96 for the normal class, respectively. Similar to the validation stage, both GB and SVM demonstrated substantial improvements after optimization. In particular, SVM showed a remarkable increase in accuracy from 55% to 92%, accompanied by improvements in precision (from 0.53 to 0.92), recall (from 0.67 to 0.97), and F1-score (from 0.59 to 0.95). GB also improved significantly, with accuracy increasing from 68% to 92%, precision from 0.72 to 0.94, recall from 0.74 to 0.95, and F1-score from 0.73 to 0.95. In contrast, RF and DT revealed minimal differences after ACO optimization. For instance, RF maintained a stable accuracy of 93%, with precision, recall, and F1-score consistently around 0.93–0.95, while DT preserved an accuracy of 90% with similarly stable precision (0.90), recall (0.93), and F1-score (0.91). The lowest performance was observed in SVM without ACO, which not only reached the lowest accuracy of 55% but also showed low precision (0.53), recall (0.67), and F1-score (0.59) across all classes.

The results at the 45° orientation also revealed that ACO optimization substantially enhanced the performance of most models (Table 10). SVM obtained the highest performance with ACO, which achieved an accuracy of 94% with precision, recall, and F1-score consistently around 0.95–0.97 (for the normal class), showing a remarkable improvement compared to the baseline SVM that only reached 60% accuracy with much lower precision (0.54), recall (0.75), and F1-score (0.63). A similar improvement was also observed in GB, where the accuracy increased from 75% to 93%, with corresponding precision, recall, and F1-score rising from approximately 0.74 to 0.95. Meanwhile, k-NN also benefited from optimization, improving its accuracy from 86% to 93% with improvements in almost all metrics. In contrast, RF and DT showed minimal differences between optimized and non-optimized versions, as RF consistently achieved high performance with an accuracy of around 93-94% and balanced precision, recall, and F1-score (0.95–0.96), while DT maintained a stable accuracy of 88–89% with only marginal changes after optimization. These findings indicate that although ACO optimization can significantly boost the performance of models such as SVM, GB, and k-NN, its impact on RF and DT remains relatively limited due to their already stable performance without optimization.

At the 90° orientation, ACO optimization demonstrated substantial performance enhancements across several models (Table 11). The most significant improvement was demonstrated by SVM, whose accuracy increased from 54% to 90%, accompanied by notable improvements in precision (from 0.56 to 0.93), recall (from 0.56 to 0.94), and F1-score (from 0.56 to 0.93). This transformation elevated SVM from the weakest baseline model to one of the strongest performers after optimization. GB also demonstrated notable progress, with accuracy rising from 69% to 88% and its precision, recall, and F1-score consistently reaching approximately 0.75–0.92. Likewise, k-NN achieved meaningful improvements, increasing from 80% to 89% accuracy with balanced precision, recall, and F1-score ranging between 0.83 and 0.93. In contrast, RF and DT showed relatively stable results, with optimization exerting only minimal impact due to their already competitive baseline performance.

Finally, at the 135° orientation, ACO optimization consistently enhanced the performance of most models, with particularly remarkable improvements observed in SVM and GB (Table 12). SVM, which initially showed weak baseline results with an accuracy of only 59%, precision of 0.54, recall of 0.77, and F1-score of 0.63, improved substantially after optimization, achieving 94% accuracy alongside precision, recall, and F1-score of 0.95, 0.97, and 0.96, respectively. Similarly, GB demonstrated a notable increase in accuracy from 73% to 93%, with precision, recall, and F1-score also rising from approximately 0.72–0.82 to 0.94–0.96. k-NN achieved a moderate improvement, with accuracy increasing from 86% to 92%, and balanced precision, recall, and F1-score reaching values of around 0.92–0.96. In contrast, RF and DT displayed relatively stable outcomes, with optimization exerting minimal influence due to their already strong baseline performance. Specifically, RF consistently maintained an accuracy of 95%, with precision, recall, and F1-score values of approximately 0.96. In contrast, DT preserved an accuracy of 90% with stable precision, recall, and F1-score values of approximately 0.91. Furthermore, Tables 9–12 also present a comparison of model performance with and without ACO-based optimization across different GLCM orientations, with particular emphasis on execution time during testing. The results show that classifiers optimized with ACO-based hyperparameters did not substantially increase the overall execution time across models. Among the models,

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

the DT consistently achieved the shortest execution time (0.0006-0.0009 seconds), reflecting its simple structure and low computational cost during prediction. The k-NN also showed fast prediction time (0.11–0.15 seconds), as classification primarily involves distance computations, which remain efficient given the dataset size.

Table 9 Comparison of the model performance with and without optimization at the 0^0 GLCM orientation

Model	Accuracy	Precision			Recall			F1-score			Time (seconds)
		Benign	Malignant	Normal	Benign	Malignant	Normal	Benign	Malignant	Normal	
RF	0.93	0.93	0.92	0.93	0.89	0.94	0.95	0.91	0.93	0.94	0.031
SVM	0.55	0.53	0.57	0.53	0.36	0.62	0.67	0.43	0.60	0.59	1.32
DT	0.90	0.89	0.90	0.90	0.87	0.90	0.93	0.88	0.90	0.91	0.00060
GB	0.68	0.63	0.70	0.72	0.61	0.70	0.74	0.62	0.70	0.73	0.0069
k-NN	0.84	0.90	0.80	0.84	0.64	0.91	0.97	0.75	0.85	0.90	0.12
RF + ACO	0.93	0.93	0.92	0.93	0.88	0.95	0.95	0.91	0.93	0.94	0.039
SVM + ACO	0.92	0.92	0.92	0.92	0.86	0.94	0.97	0.89	0.93	0.95	0.63
DT + ACO	0.90	0.89	0.91	0.90	0.88	0.90	0.93	0.88	0.90	0.91	0.00060
GB + ACO	0.92	0.91	0.91	0.94	0.88	0.94	0.95	0.89	0.93	0.95	0.024
k-NN + ACO	0.95	0.92	0.95	0.96	0.94	0.94	0.96	0.93	0.95	0.96	0.11

Table 10 Comparison of the model performance with and without optimization at the 45^0 GLCM orientation

Model	Accuracy	Precision			Recall			F1-score			Time (Seconds)
		Benign	Malignant	Normal	Benign	Malignant	Normal	Benign	Malignant	Normal	
RF	0.94	0.92	0.94	0.96	0.92	0.94	0.96	0.92	0.94	0.96	0.029
SVM	0.60	0.60	0.67	0.54	0.42	0.63	0.75	0.49	0.65	0.63	1.19
DT	0.89	0.86	0.87	0.93	0.85	0.89	0.91	0.86	0.88	0.92	0.00070
GB	0.75	0.72	0.78	0.74	0.64	0.75	0.84	0.68	0.76	0.79	0.0086
k-NN	0.86	0.88	0.83	0.86	0.72	0.90	0.96	0.79	0.87	0.91	0.13
RF + ACO	0.93	0.92	0.93	0.95	0.91	0.94	0.96	0.91	0.93	0.95	0.036
SVM + ACO	0.94	0.93	0.94	0.95	0.89	0.95	0.97	0.91	0.94	0.96	0.56
DT + ACO	0.88	0.85	0.86	0.93	0.85	0.89	0.90	0.85	0.87	0.92	0.00080
GB + ACO	0.93	0.90	0.93	0.95	0.92	0.93	0.94	0.91	0.93	0.95	0.026
k-NN + ACO	0.93	0.90	0.93	0.97	0.92	0.92	0.95	0.91	0.93	0.96	0.13

Table 11 Comparison of the model performance with and without optimization at the 90^0 GLCM orientation

Model	Accuracy	Precision			Recall			F1-score			Time (Seconds)
		Benign	Malignant	Normal	Benign	Malignant	Normal	Benign	Malignant	Normal	
RF	0.89	0.89	0.87	0.90	0.82	0.91	0.94	0.86	0.89	0.92	0.028
SVM	0.54	0.51	0.54	0.56	0.39	0.67	0.56	0.45	0.60	0.56	0.91
DT	0.84	0.82	0.85	0.86	0.80	0.83	0.91	0.81	0.84	0.88	0.00060
GB	0.69	0.67	0.66	0.75	0.59	0.72	0.76	0.63	0.69	0.75	0.0087
k-NN	0.80	0.83	0.76	0.83	0.64	0.86	0.91	0.72	0.81	0.87	0.12
RF + ACO	0.89	0.91	0.87	0.90	0.83	0.90	0.95	0.87	0.89	0.92	0.030
SVM + ACO	0.90	0.87	0.90	0.93	0.85	0.91	0.94	0.86	0.91	0.93	0.66
DT + ACO	0.84	0.83	0.83	0.85	0.78	0.82	0.91	0.80	0.83	0.88	0.00090

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

Model	Accuracy	Precision			Recall			F1-score			Time (Seconds)
		Benign	Malignant	Normal	Benign	Malignant	Normal	Benign	Malignant	Normal	
GB + ACO	0.88	0.88	0.88	0.88	0.82	0.90	0.92	0.85	0.89	0.90	0.020
k-NN + ACO	0.89	0.84	0.88	0.95	0.88	0.87	0.92	0.86	0.88	0.93	0.15

Table 12 Comparison of the model performance with and without optimization at the 135^0 GLCM orientation

Model	Accuracy	Precision			Recall			F1-score			Time (Seconds)
		Benign	Malignant	Normal	Benign	Malignant	Normal	Benign	Malignant	Normal	
RF	0.95	0.93	0.95	0.96	0.94	0.95	0.96	0.93	0.95	0.96	0.026
SVM	0.59	0.61	0.66	0.54	0.41	0.60	0.77	0.49	0.63	0.63	1.16
DT	0.90	0.88	0.91	0.91	0.88	0.91	0.90	0.88	0.91	0.90	0.00060
GB	0.73	0.70	0.77	0.72	0.63	0.73	0.82	0.66	0.75	0.77	0.0086
k-NN	0.86	0.87	0.84	0.87	0.74	0.91	0.94	0.80	0.87	0.90	0.13
RF + ACO	0.95	0.94	0.95	0.96	0.93	0.96	0.96	0.93	0.95	0.96	0.022
SVM + ACO	0.94	0.94	0.93	0.95	0.89	0.96	0.97	0.91	0.94	0.96	0.66
DT + ACO	0.90	0.88	0.91	0.91	0.88	0.92	0.91	0.88	0.92	0.91	0.00060
GB + ACO	0.93	0.90	0.94	0.96	0.92	0.93	0.94	0.91	0.93	0.95	0.022
k-NN + ACO	0.92	0.87	0.92	0.96	0.92	0.91	0.92	0.89	0.92	0.94	0.12

The RF required slightly longer times (0.026–0.039 seconds) due to its ensemble nature, involving multiple decision trees, yet still maintained sub-second efficiency. In contrast, the GB recorded slightly higher times (0.0069–0.024 seconds) compared to the RF, but remained within a very efficient range, benefiting from the relatively shallow trees used during boosting. The most time-consuming classifier was the SVM with an RBF kernel, requiring 0.56–1.32 seconds for prediction. The longer time required for processing to make predictions is explained by the quadratic optimization and kernel transformation involved in the decision function. Nevertheless, when optimized with ACO, SVM achieved improvements in accuracy, precision, recall, and F1-score across orientations.

IV. Discussion

Based on the results presented in Tables 9–12, the highest accuracy was observed at 0° and 135°, where optimized k-NN and RF achieved an accuracy of 0.95. This demonstrates that the proposed method for ultrasound image-based breast cancer classification can achieve performance levels comparable to various deep learning approaches (Table 13). Unlike deep learning models that require substantial computational resources and long training times, the proposed approach remains efficient without compromising accuracy.

Using the same dataset (BUSI), a comparison between the results of this study and those reported in previous research demonstrates the consistently better performance of the proposed method (Table 13). In contrast to the present study, which employed machine learning algorithms, Pacal [9] evaluated eleven CNN architectures for breast cancer classification and identified three models with the highest accuracies, namely VGG16 (0.854), ResNet101 (0.847), and Vision Transformer (0.886). Furthermore, Surya et al. [10] achieved an accuracy of 0.82 using MobileNetV2 optimized with Keras Tuner, while Gheflati and Rivaz [48] reported accuracies of 0.82 with the Vision Transformer and 0.83 with ResNet. In addition, Moon et al. [49] developed a CNN model based on DenseNet-161 integrated with ensemble learning, achieving an accuracy of 0.9462, with results closely approaching those obtained in this study.

Other studies, such as that by Deb and Jha [50], demonstrated that a fuzzy-rank ensemble could attain an accuracy of 0.8523, similar to the accuracy reported by Putri [51] using a conventional CNN. Meanwhile, compared with other machine learning studies using similar datasets, Yadav et al. [52] applied Topological Data Analysis for feature extraction and achieved an accuracy of only 0.684, whereas Aprilia [3], who also used GLCM for feature extraction as in this study, achieved an accuracy of 0.79. Therefore, the method proposed in this study outperformed previous studies.

According to Tables 9–12, hyperparameter optimization using ACO contributes substantially to performance improvement, particularly for SVM, which showed not only an increase in accuracy but also significant gains in recall, precision, and F1-score. The performance of SVM is highly dependent on the selection of parameters such as C, kernel type, and γ, which control the margin of the hyperplane. Consequently, when the parameters are not optimally tuned, the performance of

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

SVM tends to degrade [53]. In contrast, other algorithms remain relatively stable or experience only slight improvements, since their default performance is less sensitive to parameter tuning.

Furthermore, the comparison of model performance in Tables 9–12 shows that variations in GLCM orientation (0°, 45°, 90°, 135°) do not have a significant impact on the results. This is mainly because the second-order statistics used in GLCM are derived from the average distribution of pixel relationships. Therefore, although the calculations are based on different orientations, the resulting statistical values remain relatively consistent across angles.

Furthermore, as shown in Tables 9–12, the prediction times of the models on the testing data did not differ significantly. These differences arise from the distinct computational mechanisms of each classifier. In comparison with other studies on hyperparameter tuning using metaheuristic approaches, Ali et al. [54] examined the computational complexity of ACO, GA, Whale Optimization (WO), and PSO for optimizing SVM hyperparameters. Their findings revealed that ACO and GA achieved superior execution time and model performance compared with the other optimization techniques. However, the study by Ali et al. [54] was limited to SVM hyperparameter tuning, which, as illustrated in Fig. 3, takes longer than RF, DT, and k-NN.

Table 13 Comparison of the best result with other studies using a similar dataset

Studies	Accuracy
Aprilia et al. [3] - GLCM SVM	0.79
Pacal [9] - VGG16	0.854
Pacal [9] - ResNet101	0.847
Pacal [9] - Vision Transformer	0.886
Surya et al. [10] - MobileNetV2	0.82
Gheflati and Rivaz [48] - Vision Transformer	0.82
Gheflati and Rivaz [48] - ResNet Transformer	0.83
Moon et al. [49] - DenseNet-161	0.9462
Deb and Jha [50] - Fuzzy Rank	0.8523
Putri [51] - CNN	0.85
Yadav et al. [52] - Topological Data Analysis	0.684
Proposed Method (best model) RF + ACO (135^0) and k-NN+ ACO (0^0)	0.95

However, this study has several limitations. First, the selection of the initial and final values of the hyperparameters, as well as the number of hyperparameters used, must be determined intuitively or based on existing best practices. As a result, the more hyperparameters that need to be tested, the longer the optimization process becomes. Therefore, this study also compares the optimization time required for each classifier, showing that RF, DT, and k-NN are relatively fast in hyperparameter tuning while still achieving competitive classification performance.

The findings of this study have several important implications for the development of breast cancer detection systems based on machine learning. First, optimization using ACO not only improved model accuracy but also reduced false positives and false negatives, particularly in the SVM model. As a result, the misclassification of breast cancer by machine learning algorithms was minimized. Second, machine learning models with limited computational resources were able to achieve performance comparable to deep learning-based classification models. As shown in previous studies and this research, the computational time required for model development using machine learning was considerably lower than that of deep learning models, especially for RF, DT, and k-NN. This makes such models suitable for implementation in medical diagnostic systems operating in computationally constrained environments.

V. Conclusion

The purpose of this study was to optimize machine learning algorithms using a limited number of features, combined with ACO, aiming to achieve performance competitive with that of deep learning approaches. The findings demonstrated that applying ACO for hyperparameter tuning generally improved classification performance across all orientations. Additionally, the results indicated that variations in orientation had little effect on model performance. Classification using five GLCM-extracted features optimized with ACO outperformed both non-optimized approaches and deep learning methods previously applied by other researchers. Using the test data, the highest accuracy for each orientation was achieved with k-NN+ ACO (0.95 at 0^0), SVM+ACO (0.94 at 45^0), SVM+ACO (0.90 at 90^0), and RF+ACO (0.95 at 135^0). Although the classifiers showed improved performance after optimization, this study was limited to three classes. Further investigations are required to assess their potential for detecting breast cancer subtypes. While the present work focused on enhancing model performance using a limited set of features through optimization, future research could explore the integration of metaheuristic optimization with search-based or Bayesian approaches to achieve more robust and generalizable improvements.

References

[1] Globocan, “Global Cancer Statistics in the World in 2022,” 2022. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf

[2] American Cancer Society, “Breast Cancer Facts and Figures 2019–2020,” American Cancer Society, Inc, Atlanta, 2019.

[3] N. Aprilia and R. Rumini, “Breast Cancer Classification Based on Ultrasound Images using the Support Vector Machine (SVM) Algorithm,” *SISTEMASI*, vol. 13, no. 4, p. 1438, July 2024, doi: 10.32520/stmsi.v13i4.4113.

[4] M. Minnoor and V. Baths, “Diagnosis of Breast Cancer using Random Forests,” *Procedia Computer*

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

Science, vol. 218, pp. 429–437, 2023, doi: 10.1016/j.procs.2023.01.025.
[5] K. M. M. Uddin, N. Biswas, S. T. Rikta, and S. K. Dey, "Machine Learning-Based Diagnosis of Breast Cancer Utilizing Feature Optimization Technique," *Computer Methods and Programs in Biomedicine Update*, vol. 3, p. 100098, 2023, doi: 10.1016/j.cmpbup.2023.100098.
[6] S. Aamir *et al.*, "Predicting Breast Cancer Leveraging Supervised Machine Learning Techniques," *Computational and Mathematical Methods in Medicine*, vol. 2022, pp. 1–13, Aug. 2022, doi: 10.1155/2022/5869529.
[7] F. J. Kaunang, B. Hakim, F. Fraderic, S. Hartono, and A. K. Mulyanto, "Breast Cancer Detection using Decision Tree and Random Forest," vol. 9, no. 2, 2025, doi: 10.30871/jaic.v9i2.9073.
[8] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of Breast Ultrasound Images," *Data in Brief*, vol. 28, p. 104863, Feb. 2020, doi: 10.1016/j.dib.2019.104863.
[9] İ. Pacal, "Deep Learning Approaches for Classification of Breast Cancer in Ultrasound (US) Images," *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, vol. 12, no. 4, pp. 1917–1927, Dec. 2022, doi: 10.21597/jist.1183679.
[10] A. Surya, A. K. Shah, and J. Kabore, "Enhanced Breast Cancer Tumor Classification using MobileNetV2: A Detailed Exploration on Image Intensity, Error Mitigation, and Streamlit-driven Real-time Deployment," *ArXiv*, 2024, doi: 10.48550/arXiv.2312.03020.
[11] C. Cruz-Ramos, O. García-Avila, J.-A. Almaraz-Damian, V. Ponomaryov, R. Reyes-Reyes, and S. Sadovnychiy, "Benign and Malignant Breast Tumor Classification in Ultrasound and Mammography Images via Fusion of Deep Learning and Handcraft Features," *Entropy*, vol. 25, no. 7, p. 991, June 2023, doi: 10.3390/e25070991.
[12] M. Istighosah, A. Sunyoto, and T. Hidayat, "Breast Cancer Detection in Histopathology Images using ResNet101 Architecture," *SinkrOn*, vol. 8, no. 4, pp. 2138–2149, Oct. 2023, doi: 10.33395/sinkron.v8i4.12948.
[13] M. Korkmaz and K. Kaplan, "Effectiveness Analysis of Deep Learning Methods for Breast Cancer Diagnosis Based on Histopathology Images," *Applied Sciences*, vol. 15, no. 3, p. 1005, Jan. 2025, doi: 10.3390/app15031005.
[14] W. He, T. Zhou, Y. Xiang, Y. Lin, and J. Hu, "Deep Learning in Image Classification: Evaluating VGG19's Performance on Complex Visual Data," *ArXiv*, 2024, doi: 10.48550/arXiv.2412.20345.
[15] Y. Yanzheng, "Deep Learning Approaches for Image Classification," in *Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen, China, 2022. doi: 10.1145/3573428.3573691.
[16] R. Koulali, H. Zaidani, and M. Zaim, "Image Classification Approach using Machine Learning and an Industrial Hadoop Based Data Pipeline," *Big Data Research*, vol. 24, p. 100184, May 2021, doi: 10.1016/j.bdr.2021.100184.
[17] P. Wang, E. Fan, and P. Wang, "Comparative Analysis of Image Classification Algorithms Based on Traditional Machine Learning and Deep Learning," *Pattern Recognition Letters*, vol. 141, pp. 61–67, Jan. 2021, doi: 10.1016/j.patrec.2020.07.042.
[18] M. F. Naufal and S. F. Kusuma, "Analisis Perbandingan Algoritma Machine Learning dan Deep Learning untuk Klasifikasi Citra Sistem Isyarat Bahasa Indonesia (SIBI)," *JTIIK*, vol. 10, no. 4, pp. 873–882, Aug. 2023, doi: 10.25126/jtiik.20241046823.
[19] M. Gandor and W. Książek, "GLCM and Genetic Algorithms for Automated Diabetic Retinopathy Prediction Based on Retinal Images," presented at The Genetic and Evolutionary Computation Conference Companion (GECCO '25 Companion), New York, USA: Association for Computing Machinery, 2025. doi: https://doi.org/10.1145/3712255.3726751.
[20] A. Z. Foeady, D. C. R. Novitasari, A. H. Asyhar, and M. Firmansjah, "Automated Diagnosis System of Diabetic Retinopathy using GLCM Method and SVM Classifier," presented at the 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Malang, Indonesia: IEEE, 2018. doi: 10.1109/EECSI.2018.8752726.
[21] K. K. Mujeeb Rahman, M. Nasor, and A. Imran, "Automatic Screening of Diabetic Retinopathy using Fundus Images and Machine Learning Algorithms," *Diagnostics*, vol. 12, no. 9, p. 2262, Sept. 2022, doi: 10.3390/diagnostics12092262.
[22] L. Li *et al.*, "Enhancing Lung Cancer Detection through Hybrid Features and Machine Learning Hyperparameters Optimization Techniques," *Heliyon*, vol. 10, no. 4, p. e26192, Feb. 2024, doi: 10.1016/j.heliyon.2024.e26192.
[23] S. A. Althubiti, S. Paul, R. Mohanty, S. N. Mohanty, F. Alenezi, and K. Polat, "Ensemble Learning Framework with GLCM Texture Extraction for Early Detection of Lung Cancer on CT Images," *Computational and Mathematical Methods in Medicine*, vol. 2022, pp. 1–14, June 2022, doi: 10.1155/2022/2733965.
[24] M. Thohir, A. Z. Foeady, D. C. R. Novitasari, A. Z. Arifin, B. Y. Phiadelvira, and A. H. Asyhar, "Classification of Colposcopy Data using GLCM-SVM on Cervical Cancer," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Fukuoka, Japan: IEEE, Feb. 2020, pp. 373–378. doi: 10.1109/ICAIIC48513.2020.9065027.
[25] S. Zhang *et al.*, "MM-GLCM-CNN: A Multi-Scale and Multi-Level Based GLCM-CNN for Polyp Classification," *Computerized Medical Imaging and Graphics*, vol. 108, p. 102257, Sept. 2023, doi: 10.1016/j.compmedimag.2023.102257.

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

[26] B. E. Park, W. S. Jang, and S. K. Yoo, "Texture Analysis of Supraspinatus Ultrasound Image for Computer Aided Diagnostic System," *Healthc Inform Res*, vol. 22, no. 4, p. 299, 2016, doi: 10.4258/hir.2016.22.4.299.

[27] S. Vijayalakshmi, B. K. Pandey, D. Pandey, and M. E. Lelisho, "Innovative Deep Learning Classifiers for Breast Cancer Detection through Hybrid Feature Extraction Techniques," *Sci Rep*, vol. 15, no. 1, July 2025, doi: 10.1038/s41598-025-06669-4.

[28] Y. Hao *et al.*, "Breast Cancer Histopathological Images Classification Based on Deep Semantic Features and Gray Level Co-Occurrence Matrix," *PLoS ONE*, vol. 17, no. 5, p. e0267955, May 2022, doi: 10.1371/journal.pone.0267955.

[29] M. Z. Al Ghifari, W. Ferriastuti, T. Harsono, R. Sigit, and F. Hayati, "Brain Tumour Segmentation in MRI Data using Gray Level Co-Occurrence Matrix," presented at the 2024 International Conference on Electrical and Information Technology (IEIT), Malang, Indonesia: IEEE, 2024. doi: 10.1109/IEIT64341.2024.10763342.

[30] A. Fauzi and L. Evan Lubis, "Optimization of Retinal Blood Vessel Segmentation Based on Gabor Filters and Particle Swarm Optimization," *IJEECS*, vol. 29, no. 3, p. 1590, Mar. 2023, doi: 10.11591/ijeecs.v29.i3.pp1590-1596.

[31] G. Lin, J. Jiang, J. Bai, Y. Su, Z. Su, and H. Liu, "Frontiers and Developments of Data Augmentation for Image: From Unlearnable to Learnable," *Information Fusion*, vol. 114, p. 102660, Feb. 2025, doi: 10.1016/j.inffus.2024.102660.

[32] P. Ramamoorthy, B. R. Ramakantha Reddy, S. S. Askar, and M. Abouhawwash, "Histopathology-Based Breast Cancer Prediction using Deep Learning Methods for Healthcare Applications," *Front. Oncol.*, vol. 14, p. 1300997, June 2024, doi: 10.3389/fonc.2024.1300997.

[33] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.

[34] O. P. Andra, "Optimasi Gray Level Co-Occurance Matrix (GLCM) Menggunakan Metode Ant Colony Optimization (ACO) pada Klasfikasi Pengenalan Daging Sapi dan Daging Babi," *UIN Sultan Syarif Kasim Riau*, 2020.

[35] T. Chen, C. Yang, L. Han, and S. Guo, "GF-2 Data for Lithological Classification using Texture Features and PCA/ICA Methods in Jixi, Heilongjiang, China," *Remote Sensing*, vol. 15, no. 19, p. 4676, Sept. 2023, doi: 10.3390/rs15194676.

[36] W. Chen, K. Yang, Z. Yu, Y. Shi, and C. L. P. Chen, "A Survey on Imbalanced Learning: Latest Research, Applications and Future Directions," *Artif Intell Rev*, vol. 57, no. 6, p. 137, May 2024, doi: 10.1007/s10462-024-10759-6.

[37] S. H. Hasanah, "Classification Support Vector Machine In Breast Cancer Patients," *BAREKENG: J. Il. Mat. & Ter.*, vol. 16, no. 1, pp. 129–136, Mar. 2022, doi: 10.30598/barekengvol16iss1pp129-136.

[38] Asri Mulyani, Sarah Khoerunisa, and Dede Kurniadi, "Perbandingan Kinerja Algoritma KNN dan SVM Menggunakan SMOTE untuk Klasifikasi Penyakit Diabetes," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 14, no. 1, pp. 25–34, Feb. 2025, doi: 10.22146/jnteti.v14i1.15198.

[39] M. S. Tahosin, M. A. Sheakh, T. Islam, R. J. Lima, and M. Begum, "Optimizing Brain Tumor Classification through Feature Selection and Hyperparameter Tuning in Machine Learning Models," *Informatics in Medicine Unlocked*, vol. 43, p. 101414, 2023, doi: 10.1016/j.imu.2023.101414.

[40] C. Kavitha, V. Mani, S. R. Srividhya, O. I. Khalaf, and C. A. Tavera Romero, "Early-Stage Alzheimer's Disease Prediction using Machine Learning Models," *Front. Public Health*, vol. 10, p. 853294, Mar. 2022, doi: 10.3389/fpubh.2022.853294.

[41] A. Hemmati-Sarapardeh, A. Larestani, M. Nait Amar, and S. Hajirezaie, "Chapter 3—Training and Optimization Algorithms.," in *Applications of Artificial Intelligence Techniques in the Petroleum Industry*, Gulf Professional Publishing, 2020, pp. 51–78. [Online]. Available: https://doi.org/10.1016/C2018-0-04421-7

[42] L. Phan-Van, H. Takano, and T. Nguyen Duc, "A Comparison of Different Metaheuristic Optimization Algorithms on Hydrogen Storage-Based Microgrid Sizing," *Energy Reports*, vol. 9, pp. 542–549, Oct. 2023, doi: 10.1016/j.egyr.2023.05.152.

[43] S. Akbari, H. Ramazi, and R. Ghezelbash, "A Novel Framework for Optimizing the Prediction of Areas Favorable to Porphyry-Cu Mineralization: Combination of Ant Colony and Grid Search Optimization Algorithms with Support Vector Machines," *Nat Resour Res*, vol. 34, pp. 703–729, doi: 10.1007/s11053-024-10431-4.

[44] L. Ben Said *et al.*, "Harnessing Meta-Heuristic, Bayesian, and Search-Based Techniques in Optimizing Machine Learning Models for Improved Energy Storage with Microencapsulated PCMs," *International Communications in Heat and Mass Transfer*, vol. 162, p. 108537, Mar. 2025, doi: 10.1016/j.icheatmasstransfer.2024.108537.

[45] A. Gjecka and M. Fetaji, "A Comparative Study of Thyroid Data Classification Based on GA, BPSO, and ACO Metaheuristics Approaches," presented at the International Conference on Mathematical and Statistical Physics, Computational Science, Education and Communication (ICMSCE 2023), 2023. doi: 10.1117/12.3011406.

[46] A. Saif Alghawli and A. I. Taloba, "An Enhanced Ant Colony Optimization Mechanism for the Classification of Depressive Disorders," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–12, June 2022, doi: 10.1155/2022/1332664.

[47] A. M. Adrian, A. Utamima, and K.-J. Wang, "A Comparative Study of GA, PSO and ACO for Solving Construction Site Layout Optimization,"

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

KSCE Journal of Civil Engineering, vol. 19, no. 3, pp. 520–527, Mar. 2015, doi: 10.1007/s12205-013-1467-6.
[48] B. Gheflati and H. Rivaz, "Vision Transformer for Classification of Breast Ultrasound Images," Feb. 12, 2025, *arXiv*: arXiv:2110.14731. doi: 10.48550/arXiv.2110.14731.
[49] W. K. Moon, Y.-W. Lee, H.-H. Ke, S. H. Lee, C.-S. Huang, and R.-F. Chang, "Computer-Aided Diagnosis of Breast Ultrasound Images using Ensemble Learning from Convolutional Neural Networks," *Computer Methods and Programs in Biomedicine*, vol. 190, p. 105361, July 2020, doi: 10.1016/j.cmpb.2020.105361.
[50] S. D. Deb and R. K. Jha, "Breast UltraSound Image Classification using Fuzzy-Rank-Based Ensemble Network," *Biomedical Signal Processing and Control*, vol. 85, p. 104871, Aug. 2023, doi: 10.1016/j.bspc.2023.104871.
[51] D. M. Putri, M. Ikhsan, S. Nurjanah, B. W. Akramunnas, and A. Rahmawati, "A CNN-Based Approach for Breast Cancer Classification from Ultrasound Images," vol. 12, no. 1.
[52] A. Yadav, F. Nisha, and B. Coskunuzer, "Breast Cancer Detection with Topological Machine Learning," in *Proceedings of the 2023 10th International Conference on Biomedical and Bioinformatics Engineering*, Kyoto Japan: ACM, Nov. 2023, pp. 217–222. doi: 10.1145/3637732.3637744.
[53] Y.-J. Chang, Y.-L. Lin, and P.-F. Pai, "Support Vector Machines with Hyperparameter Optimization Frameworks for Classifying Mobile Phone Prices in Multi-Class," *Electronics*, vol. 14, no. 11, p. 2173, May 2025, doi: 10.3390/electronics14112173.
[54] Y. Ali, E. Awwad, M. Al-Razgan, and A. Maarouf, "Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity," *Processes*, vol. 11, no. 2, p. 349, Jan. 2023, doi: 10.3390/pr11020349.

Author Biography

Ahmad Fauzi is a lecturer in the Information Systems Study Program, UIN Imam Bonjol Padang. He earned his bachelor's degree in Physics from Universitas Indonesia in 2012 and his master's degree in Information Technology from the same university in 2017. In addition to his academic career, he is actively engaged as an information technology practitioner in several private and government institutions, contributing to the development and implementation of digital systems and data-driven solutions. His research interests encompass information systems, computer vision, and the application of computational and artificial intelligence approaches in interdisciplinary fields, particularly those involving the integration of technology to enhance decision-making and organizational performance.

Lukmanda Evan Lubis received the degrees of B.Sc., M.Sc., and Ph.D. in physics, with a specialization in medical physics, from the Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia in 2011, 2014, and 2021, respectively. Since 2016, he has been appointed as a lecturer in the same department. With his area of expertise being diagnostic and interventional radiology physics, he also serves as a consultant medical physicist at Universitas Indonesia Hospital. He is also a co-founder and Manager of Education and Training at the Center for Medical Physics and Biophysics, Universitas Indonesia. Outside the campus, he holds positions in the Indonesian Association of Physicists in Medicine (as Deputy Secretary General in 2015, Secretary General in 2019, and Chairperson for Research and Scientific Publication in 2022) and in the South East Asian Federation of Organizations for Medical Physics (SEAFOMP) as an Executive Committee member. He is a member of the Indonesian Radiation Protection Society, an international affiliate of the American Association of Physicists in Medicine, and an individual member of the European Federation of Organizations for Medical Physics. He can be contacted at email: lukmanda.evan@sci.ui.ac.id

Raju Wandira received his bachelor's degree from Institut Agama Islam Negeri (IAIN) Bukittinggi and completed his master's degree at Universitas Putra Indonesia "YPTK" Padang. He is currently serving as a lecturer at a university in Padang, where he is actively involved in teaching, research, and academic development. His primary research interests are in the field of information systems, with a particular focus on leveraging technology to enhance data management, decision-making, and business process optimization. In addition to his academic roles, he participates in various research collaborations and community service programs that emphasize the practical implementation of information systems within educational institutions and organizational environments to improve operational efficiency and innovation.

Syarto Musthofa is a lecturer in the Mathematics Program at the Faculty of Science and Technology, UIN Imam Bonjol Padang. He obtained his Bachelor's degree in Mathematics from the Faculty of Mathematics and Natural Sciences at the State University of Medan in 2015 and his Master of Science degree in Mathematics from Gadjah Mada University, Yogyakarta, in 2019. His academic and research interests

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266

focus on computational statistics and biostatistics, with a particular emphasis on survival analysis, longitudinal data analysis, and high-dimensional data modeling. In addition to teaching and research, he is actively involved in academic collaborations and statistical consulting, especially in projects related to applied mathematics and data analysis for health and social sciences.

Corresponding author: Ahmad Fauzi, ahmadfauzi@uinib.ac.id, Information Systems Study Program, UIN Imam Bonjol Padang, Padang, Indonesia.
Digital Object Idenfier (DOI): https://doi.org/10.35882/ijeeemi.v7i4.266