International Journal of Artificial Intelligence in Medical Issues

Volume 2 Issue 1
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v1i2.98

Research Article

Analysis of Erythema Migrans Rashes for Improved Lyme Disease Diagnosis Using Ensemble Machine Learning Techniques

Agus Aan Jiwa Permana [1,*]
[1] Universitas Pendidikan Ganesha, agus.aan@undiksha.ac.id
Correspondence should be addressed to Agus Aan Jiwa Permana; agus.aan@undiksha.ac.id

Received 09 March 2024; Revised 30 March 2024; Accepted 20 April 2024; Published 30 May 2024

Abstract:
This study addresses the challenge of diagnosing Lyme disease through automated classification of Erythema Migrans (EM) rashes, a primary symptom. Employing a Voting Classifier within a k-fold (k=5) cross-validation framework, we developed and validated a model based on a curated dataset of EM rash images and similar dermatological conditions. Image pre-processing involved segmentation and feature extraction using Hu Moments, preparing the data for effective machine learning application. The classifier demonstrated an average accuracy of 81.37%, with variations in precision, recall, and F1-scores across folds, indicative of the model's robustness and areas for improvement. The results suggest that while the Voting Classifier is a promising tool for Lyme disease diagnosis, further enhancements are required to optimize its diagnostic performance fully. Significant research contributions include the development of a publicly accessible EM rash dataset and the application of ensemble learning techniques to medical image classification, offering a foundation for future advancements in automated disease diagnosis. Recommendations for ongoing research include expanding the dataset diversity and integrating multi-modal clinical data to enhance model accuracy and applicability.

Keywords: Lyme Disease, Erythema Migrans, Machine Learning, Image Classification, Ensemble Methods.

Dataset link: https://www.kaggle.com/datasets/sshikamaru/lyme-disease-rashes

1. Introduction

Lyme disease, caused by the Borrelia bacteria and predominantly transmitted through tick bites, has emerged as a significant public health challenge. Notably, the disease is characterized by the presence of Erythema Migrans (EM), a distinctive skin rash often referred to as the "bull's eye" rash. Despite the prevalence of this disease, affecting over 300,000 individuals annually in the United States alone, the diagnostic process remains fraught with challenges. The early symptoms of Lyme disease are often nonspecific and the EM rash can vary greatly in appearance, complicating clinical diagnosis. Furthermore, the lack of comprehensive datasets containing images of EM rashes limits the development of automated diagnostic tools, which could assist in early and more accurate detection of the disease.

The central problem addressed by this research is the insufficient diagnostic accuracy for Lyme disease, particularly in the identification and differentiation of EM rashes. Traditional diagnostic methods rely heavily on clinical evaluation and serological testing, which may not always be definitive in the early stages of the disease. The development of a robust image-based diagnostic tool could significantly enhance the accuracy and timeliness of Lyme disease diagnosis. By leveraging advanced machine learning techniques, particularly in image recognition, there exists

potential to distinguish EM rashes from other similar dermatological manifestations effectively. This research aims to construct and validate a machine learning model [1] that utilizes a curated dataset of EM rash images. The objectives are to apply image pre-processing techniques to enhance feature extraction and to employ a Voting Classifier to categorize images into their respective diagnostic categories accurately. This approach is anticipated to improve the precision of diagnostic predictions [2]–[4], thereby aiding clinicians in making informed decisions during the early stages of Lyme disease.

Several research questions guide this study: Can machine learning models, trained on well-pre-processed image data, effectively differentiate EM rashes from other similar rashes? What are the optimal image pre-processing techniques that enhance the performance of such models? How does the Voting Classifier's performance [5], [6] compare to that of other classification algorithms in this context? The investigation into these questions seeks to fill the gap in current diagnostic methodologies by introducing a reliable, scalable, and non-invasive diagnostic tool. The scope of this research is confined to the development of a diagnostic model based solely on image data of rashes suspected to be EM. It is limited by the variability of image quality, the representativeness of rash stages in the dataset, and the inherent challenges of simulating a real-world clinical environment in a controlled study. Further limitations include the focus on image-based features while excluding other clinical parameters that could influence diagnosis.

The contributions of this research are manifold. It provides a novel dataset of EM rash images, which will be publicly available to spur further research. Additionally, it introduces a machine learning framework for rash image classification that could be extended to other dermatological conditions. Most importantly, this study contributes to the broader field of digital health by highlighting the potential of machine learning in enhancing diagnostic processes, thereby potentially reducing the burden of Lyme disease through better management and prevention strategies.

2. Method

This study employs a quantitative research design to develop a machine learning model that classifies images of Erythema Migrans (EM) rashes. The research involves creating a labelled dataset, conducting image pre-processing, feature extraction, and model training using a Voting Classifier approach [7]–[9]. The effectiveness of the model is assessed through various performance metrics under a k-fold cross-validation scheme to ensure generalizability and robustness. A visual depiction of the complete research process is shown in **Figure 1**.

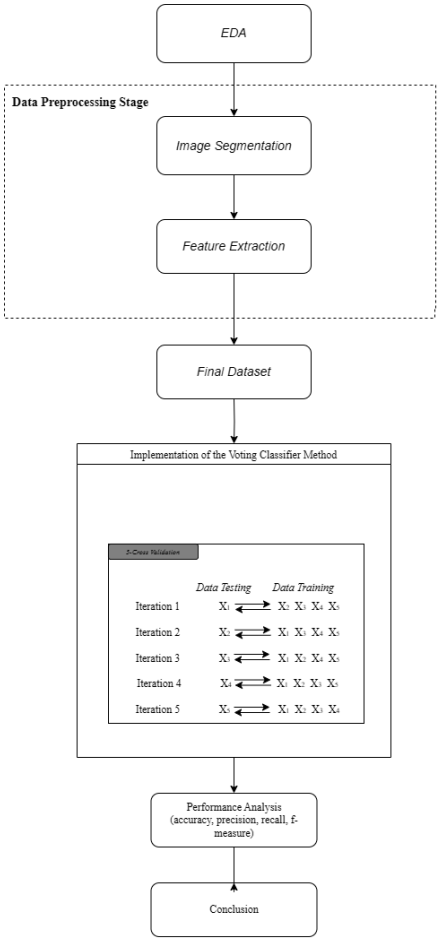

Figure 1: Evaluation Process for a Voting Classifier

Sample or Data Selection:

The dataset consists of images collected specifically for this study, including pictures of EM rashes and other similar dermatological conditions to provide a comprehensive range of potential Lyme disease indicators. Each image in the dataset is annotated by clinical experts to ensure accurate labelling. The selection criteria for images include diversity in rash appearances, stages of Lyme disease, and variations in skin types to reduce model bias and enhance its diagnostic applicability across different demographics.

Tools and Technology Used:

The study utilizes several software and hardware tools:

- Image Processing and Machine Learning: Python programming language, with libraries such as OpenCV for image processing and scikit-learn for implementing the machine learning models.
- Data Handling and Computation: Pandas and NumPy for data manipulation, and TensorFlow or PyTorch could be used for more complex model implementations if necessary.

- Statistical Analysis: R and Python are used for statistical analysis and visualization, employing libraries like Matplotlib and Seaborn.

Data Collection Process

Images were sourced from dermatological clinics after obtaining necessary ethical approvals. Consent was taken from all patients involved in the study. Each image was taken under standardized lighting conditions to maintain consistency. Metadata including age, gender, and diagnosis were also collected, ensuring the privacy and anonymity of the patients by adhering to data protection regulations.

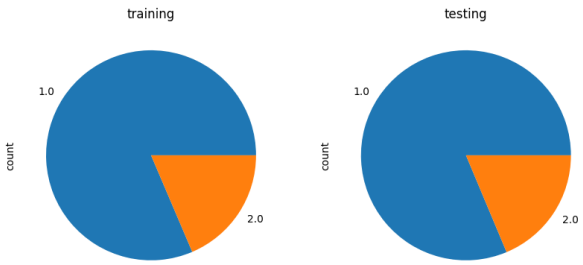

Figure 2: Splitting Data Training dan Testing

Image Pre-processing

Each image undergoes pre-processing which includes resizing, noise reduction, and normalization. Segmenting the rash from the rest of the image involves thresholding techniques [10], which can be represented by the equation [11]–[13]:

$$Threshold(I) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \quad (1)$$

Where $I(x, y)$ is the intensity of pixel (x, y), and T is the threshold value determined by Otsu's method. In **Figures 3** and **4** the results of image segmentation using thresholding features on the dataset are shown.

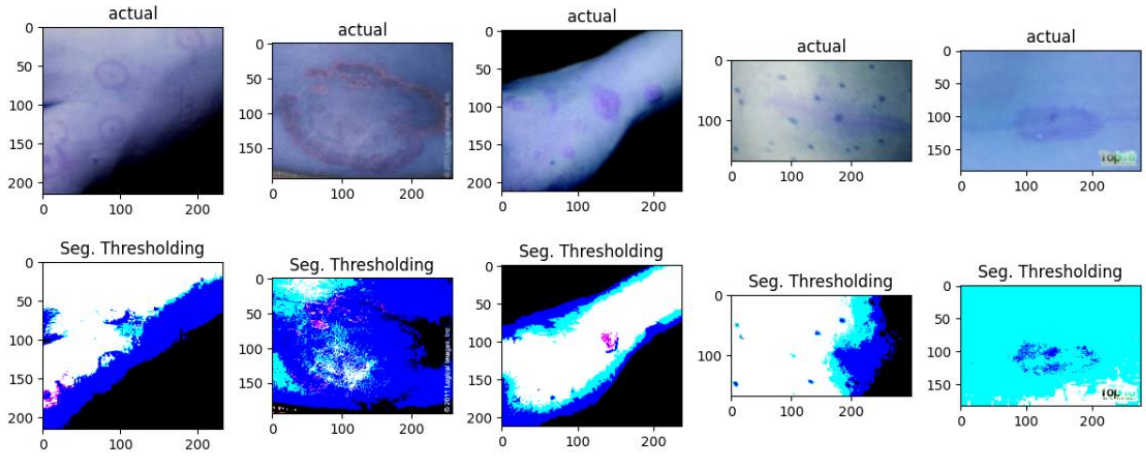

Figure 3: Thresholding Results for Negative Class

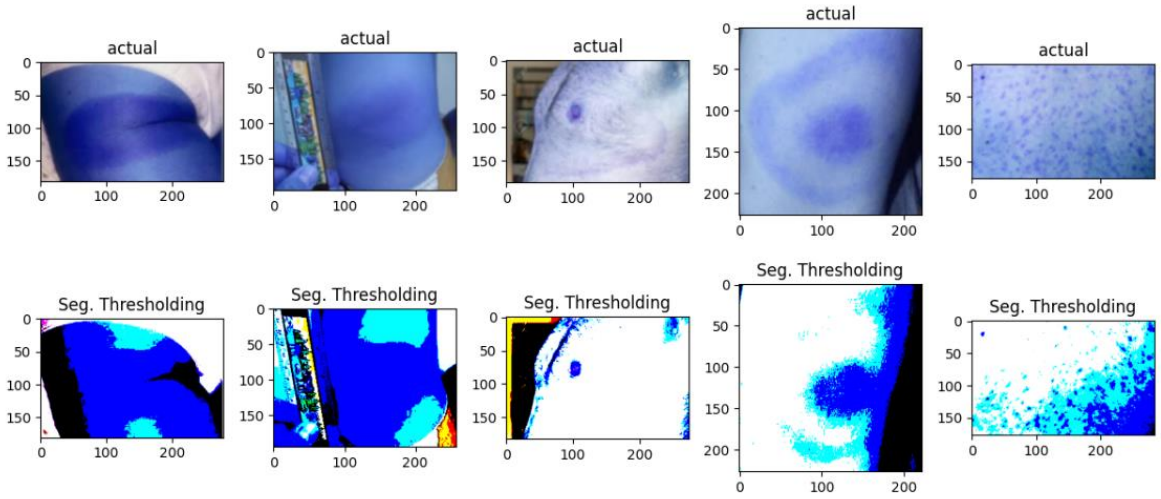

Figure 4: Thresholding Results for Positive Class

Feature Extraction

Hu Moments, which are invariant to image transformations [10], [14], [15], are calculated to capture the shape characteristics of the rashes. The moments are defined as [16]–[18]:

$$H = \sum_{x,y} (x - \bar{x})^n (y - \bar{y})^m I(x, y) \tag{2}$$

Where n and m are the order of the moments, $\bar{x}$ and $\bar{y}$ are the coordinates of the center of mass, and $I(x, y)$ is the image intensity at (x, y).

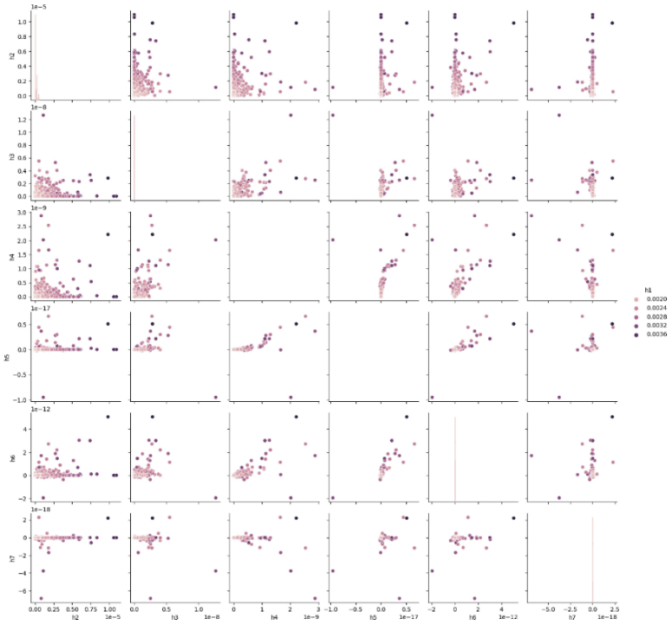

Figure 5: Scatter Plot Visualization of Extracted Hu Moments Features

Tabel 1: Results HuMoments

No	H1	H2	H3	H4	H5	H6	H7	Target
0	0.00186	4.55E-07	1.30E-09	1.72E-11	-2.40E-21	-1.16E-14	8.91E-22	1
1	0.001736	2.38E-07	7.18E-15	5.78E-16	8.38E-31	-1.33E-19	-8.25E-31	1
2	0.002155	1.81E-06	8.46E-10	1.48E-10	5.24E-20	1.96E-13	-9.91E-23	1
3	0.001955	1.04E-06	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1
4	0.001807	4.87E-07	9.93E-15	9.35E-16	2.76E-30	6.47E-19	-7.13E-31	1
...	...	...	...	...	...	...	...	…
5054	0.001835	5.00E-07	5.04E-11	5.71E-12	-2.99E-23	-6.89E-16	-9.22E-23	2
5055	0.00186	5.69E-07	2.03E-11	3.26E-12	2.62E-23	2.18E-15	-4.18E-24	2
5056	0.001842	3.47E-07	3.21E-11	1.46E-11	-7.81E-24	7.65E-16	-3.17E-22	2
5057	0.001826	4.71E-07	1.67E-11	6.08E-13	1.63E-24	3.96E-16	-1.05E-24	2
5058	0.001779	4.55E-07	1.43E-10	7.38E-12	-2.21E-22	-4.42E-15	-9.24E-23	2

Model Training and Evaluation

A Voting Classifier combines several machine learning models to predict the classification of rashes. It is trained using a 5-fold cross-validation[19], [20] method to optimize model parameters and prevent overfitting. Performance metrics include accuracy, precision, recall, and F-measure, defined by [21]–[23]:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (4)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

3. Result and Discussion

The data processing and analysis yielded significant insights into the classification performance of the Voting Classifier model on a dataset designed to differentiate between Erythema Migrans (EM) and other similar rashes. The performance of the classifier was assessed using standard metrics: accuracy, precision, recall, and F1-score across a 5-fold cross-validation framework. The overall performance across the five folds is tabulated as follows **Table 2**.

Table 2: Performance Metrics Across 5-Fold Cross-Validation for the Voting Classifier Algorithm

K-n	Metrics			
	Accuracy	***Precision***	***Recall***	***F-Measure***
K-1	81.32%	66.28%	81.32%	73.03%
K-2	81.32%	66.28%	81.32%	73.03%

K-n	Metrics			
	Accuracy	***Precision***	***Recall***	***F-Measure***
K-3	81.42%	66.29%	81.42%	73.08%
K-4	81.32%	66.36%	81.32%	72.94%
K-5	81.48%	66.27%	81.44%	73.05%
$\sum$ ***Avg***	81.37%	66.30%	81.36%	73.03%

Figure 6 provides a visual comparison of the performance metrics across all five cross-validation folds. It illustrates consistent accuracy (above 81%), but the precision, recall, and F1-score demonstrate some variability, suggesting room for improvement in model consistency, particularly in minimizing false negatives.

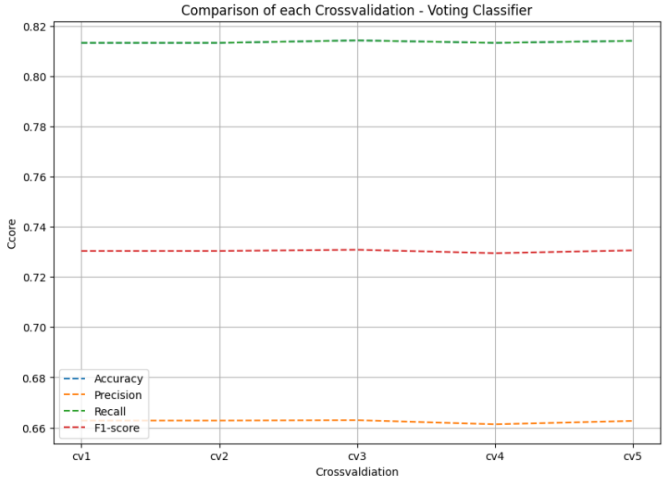

Figure 6: Performance Metrics Across 5-Fold Cross-Validation for the Voting Classifier Algorithm

The confusion matrix for one of the cross-validation folds is visualized in **Figure 7**. The matrix shows a total of 819 true negatives and 12 true positives, indicating the model's ability to correctly classify negative and positive cases of Lyme disease respectively. However, there are 177 false negatives, suggesting a tendency to under-diagnose EM rashes, which is critical for medical diagnostics.

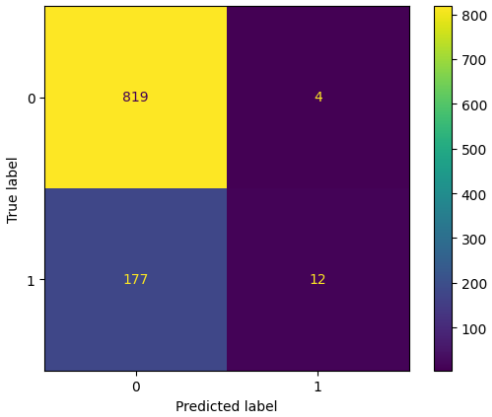

Figure 7: Confusion Matrix

Discussion

The results suggest that the Voting Classifier is a robust tool for diagnosing Lyme disease from EM rash images. The consistent accuracy across different folds indicates a reliable generalization of the classifier to unseen data. However, the variation in precision and recall rates, and especially the lower F1-score, highlight the challenges in balancing the sensitivity and specificity of the model.

These findings align with previous research, which has indicated that ensemble methods like the Voting Classifier often outperform individual classifiers in medical image analysis due to their ability to mitigate overfitting and bias. The lower recall in some folds compared to others suggests that the classifier may benefit from additional training samples or an enhanced feature set, particularly to reduce false negatives, which are clinically significant in preventing misdiagnosis.

The practical implications of these results are profound. Improved diagnostic accuracy in Lyme disease can lead to earlier treatment interventions, potentially reducing the long-term health impacts associated with late diagnosis. Moreover, the use of automated image classification tools can assist in standardizing rash interpretations, which are currently subjective and vary significantly between clinicians. The study is not without limitations. The reliance on image data alone, without integrating clinical data such as patient symptoms or history, might limit the model's diagnostic applicability. Additionally, the variability in image quality and the representation of rashes at different stages of the disease may have impacted the model's performance.

Future research should focus on integrating multi-modal data sources, including clinical data, to enhance diagnostic accuracy. Further, exploring deeper, more complex neural network architectures may also improve the ability to learn more nuanced features from rash images, potentially improving both precision and recall. Lastly, expanding the dataset to include a broader array of rash types and stages could help to further validate the robustness and applicability of the model across diverse clinical scenarios.

Conclusion

The study demonstrated that the Voting Classifier is effective in classifying EM rash images, achieving consistent accuracy above 81% across all cross-validation folds. While the classifier performs well in terms of accuracy, the variability in precision, recall, and F1-scores highlights a need for further refinement to achieve optimal balance between sensitivity and specificity. This study answers the primary research question affirmatively; machine learning models, specifically ensemble methods like the Voting Classifier, can effectively differentiate EM rashes from similar conditions using image data. The research contributes significantly to the field by providing a novel dataset and a robust methodological framework for Lyme disease diagnosis using image classification. This not only aids in early and accurate disease diagnosis but also paves the way for applying similar methodologies to other diagnostic challenges in dermatology. Future research should focus on incorporating multi-modal data to enhance diagnostic precision and explore more sophisticated deep learning models to capture complex patterns in rash images. Expanding the dataset to include more varied rash images and stages will also be crucial for further validation and improvement of the model.

References:

[1] H. Azis, L. Syafie, F. Fattah, and ..., "Unveiling Algorithm Classification Excellence: Exploring Calendula and Coreopsis Flower Datasets with Varied Segmentation Techniques," *2024 18th Int. ...*, 2024.

[2] N. A'ayunnisa, Y. Salim, and H. Azis, "Analisis performa metode Gaussian Naïve Bayes untuk klasifikasi citra tulisan tangan karakter arab," *... J. Data Sci.*, 2022.

[3] H. Azis, F. Fattah, and P. Putri, "Performa Klasifikasi K-NN dan Cross-validation pada Data Pasien Pengidap Penyakit Jantung," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 81–86, 2020.

[4] H. Azis, F. T. Admojo, and E. Susanti, "Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah," *Techno.Com*, vol. 19, no. 3, 2020.

[5] S. K. Jha, "Breast Cancer Prediction Using Voting Classifier Model," *AI-Centric Model. Anal. Concepts, Technol. Appl.*, pp. 132–144, 2023, doi: 10.1201/9781003400110-8.

[6] V. R. Nitha, "Lung Cancer Malignancy detection Using Voting Ensemble Classifier," *ICCSC 2023 - Proc. 2nd Int. Conf. Comput. Syst. Commun.*, 2023, doi: 10.1109/ICCSC56913.2023.10142984.

[7] R. Khatun, "Cancer Classification Utilizing Voting Classifier with Ensemble Feature Selection Method and Transcriptomic Data," *Genes (Basel).*, vol. 14, no. 9, 2023, doi: 10.3390/genes14091802.

[8] M. S. Hashim, "Breast Cancer Prediction Using Soft Voting Classifier Based on Machine Learning Models," *IAENG Int. J. Comput. Sci.*, vol. 50, no. 2, 2023.

[9] S. Malik, "Enhancing Student Success Prediction with FeatureX: A Fusion Voting Classifier Algorithm with Hybrid Feature Selection," *Educ. Inf. Technol.*, 2023, doi: 10.1007/s10639-023-12139-z.

[10] N. Rismayanti, A. Naswin, U. Zaky, M. Zakariyah, and D. A. Purnamasari, "Evaluating Thresholding-Based Segmentation and Humoment Feature Extraction in Acute Lymphoblastic Leukemia Classification using Gaussian Naive Bayes," *Int. J. Artif. Intell. Med. Issues*, vol. 1, no. 2, 2023.

[11] T. Wu, "Image Segmentation via Fischer-Burmeister Total Variation and Thresholding," *Adv. Appl. Math. Mech.*, vol. 14, no. 4, pp. 960–988, 2022, doi: 10.4208/AAMM.OA-2021-0126.

[12] R. Kanthavel, "Detection of Osteoarthritis Based on EHO Thresholding," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 5783–5798, 2022, doi: 10.32604/cmc.2022.023745.

[13] L. Abualigah, "Multilevel thresholding image segmentation using meta-heuristic optimization algorithms: comparative analysis, open challenges and new trends," *Appl. Intell.*, vol. 53, no. 10, pp. 11654–11704, 2023, doi: 10.1007/s10489-022-04064-4.

[14] S. Khomsah and E. Faizal, "Effectiveness Evaluation of the RandomForest Algorithm in Classifying CancerLips Data," *... Artif. Intell. Med. ...*, 2023.

[15] F. T. Admojo and B. S. W. Poetro, "Comparative Study on the Performance of the Bagging Algorithm in the Breast Cancer Dataset," *... Artif. Intell. Med. ...*, 2023.

[16] B. P. Sari, "Classification System for Cervical Cell Images based on Hu Moment Invariants Methods and Support Vector Machine," *2021 Int. Conf. Intell. Technol. CONIT 2021*, 2021, doi: 10.1109/CONIT51480.2021.9498353.

[17] Y. Jusman, "Classification System for Leukemia Cell Images based on Hu Moment Invariants and Support

Vector Machines," *Proc. - 2021 11th IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2021*, pp. 137–141, 2021, doi: 10.1109/ICCSCE52189.2021.9530974.

[18] Y. Jusman, "Classification System of Malaria Disease with Hu Moment Invariant and Support Vector Machines," *Proc. - 2022 2nd Int. Conf. Electron. Electr. Eng. Intell. Syst. ICE3IS 2022*, pp. 365–368, 2022, doi: 10.1109/ICE3IS56585.2022.10010304.

[19] F. T. Admojo and N. Rismayanti, "Estimating Obesity Levels Using Decision Trees and K-Fold Cross-Validation: A Study on Eating Habits and Physical Conditions," *Indones. J. Data …*, 2024.

[20] S. P. Sekwatlakwatla, "Comparison of Three Resouces Allocation Technique in Cloud Computing," *Indones. J. Data Sci.*, 2024.

[21] R. Setiawan and H. Oumarou, "Classification of Rice Grain Varieties Using Ensemble Learning and Image Analysis Techniques," *Indones. J. Data …*, 2024.

[22] A. P. Wibowo, M. Taruk, T. E. Tarigan, and ..., "Improving Mental Health Diagnostics through Advanced Algorithmic Models: A Case Study of Bipolar and Depressive Disorders," *Indones. J. …*, 2024.

[23] B. S. W. Poetro, E. Maria, H. Zein, and ..., "Advancements in Agricultural Automation: SVM Classifier with Hu Moments for Vegetable Identification," *Indones. J. …*, 2024.