International Journal of Artificial Intelligence in Medical Issues

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

Research Article

Performance Metrics of AdaBoost and Random Forest in Multi-Class Eye Disease Identification: An Imbalanced Dataset Approach

Thomas Edyson Tarigan [1,*], Erma Susanti [2], M. Ikbal Siami [3], Ika Arfiani [4], Agus Aan Jiwa Permana [5], I Made Sunia Raharja [6]

[1] Universitas Teknologi Digital Indonesia, tarigan@utdi.ac id
[2] Institut Sains & Teknologi AKPRIND, erma@akprind.ac.id
[3] Institut Teknologi dan Bisnis STIKOM Ambon, siami25@gmail.com
[4] Universitas Ahmad Dahlan, ika.arfiani@tif.uad.ac.id
[5] Universitas Pendidikan Ganesha, agus.aan@undiksha.ac.id
[6] Universitas Udayana, sunia.raharja@unud.ac.id
Correspondence should be addressed to Thomas Edyson Tarigan; tarigan@utdi.ac id

Received 09 July 2023; Revised 30 July 2023; Accepted 20 September 2023; Published 30 November 2023

Abstract:
This study presents a comprehensive evaluation of AdaBoost and Random Forest Classifier algorithms in the classification of eye diseases, focusing on a challenging scenario involving an imbalanced dataset. Eye diseases, particularly Cataract, Diabetic Retinopathy, Glaucoma, and Normal eye conditions, pose significant diagnostic challenges, and the advent of machine learning offers promising avenues for enhancing diagnostic accuracy. Our research utilizes a dataset preprocessed with Canny edge detection for image segmentation and Hu Moments for feature extraction, providing a robust foundation for the comparative analysis. The performance of the algorithms is assessed using a 5-fold cross-validation approach, with accuracy, precision, recall, and F1-score as the key metrics. The results indicate that the Random Forest Classifier outperforms AdaBoost across these metrics, albeit with moderate overall performance. This finding underscores the potential and limitations of using advanced machine learning techniques for medical image analysis, particularly in the context of imbalanced datasets. The study contributes to the field by providing insights into the effectiveness of different machine learning algorithms in handling the complexities of medical image classification. For future research, it recommends exploring a diverse range of image processing techniques, delving into other sophisticated machine learning models, and extending the study to encompass a wider array of eye diseases. These findings have practical implications in guiding the selection of machine learning tools for medical diagnostics and highlight the need for continuous improvement in automated systems for enhanced patient care.

Keywords: AdaBoost, Random Forest Classifier, Eye Disease Classification, Machine Learning, Imbalanced Dataset, Medical Image Analysis, Canny Edge Detection, Hu Moments, Diagnostic Accuracy.

Dataset link: https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification/

1. Introduction

In the realm of ophthalmology, accurate diagnosis of eye diseases is paramount for effective treatment and management. Advances in machine learning and image processing have opened new frontiers in medical diagnostics, particularly in the classification of eye diseases through retinal images. Despite these technological strides, challenges persist in the accurate classification of diseases such as Cataract, Diabetic Retinopathy, Glaucoma, and Normal eye conditions. The complexity is compounded when dealing with imbalanced datasets, a common occurrence in medical imaging, where certain conditions are underrepresented compared to others. This discrepancy often leads to biased models, undermining the diagnostic accuracy critical in medical applications. Therefore, exploring robust machine learning techniques that can efficiently handle such imbalances becomes imperative.

The primary problem addressed in this research is the effective classification of eye diseases in the presence of imbalanced datasets. Traditional machine learning models tend to underperform when faced with class imbalance, leading to a higher misclassification rate of the minority class. This issue is particularly acute in the field of medical imaging, where accurate classification can be a matter of crucial medical intervention. The challenge extends to the selection of appropriate feature extraction and image segmentation techniques that can further enhance the performance of these classification models.

Our study aims to bridge this gap by exploring and comparing the effectiveness of two advanced machine learning algorithms: AdaBoost [1]–[3] and Random Forest Classifier. These algorithms were selected for their known robustness and ability to handle complex classification tasks. The research focuses on evaluating their performance in the context of an imbalanced dataset for eye disease classification. The objective is to determine which algorithm provides better accuracy, precision, recall, and F-measure [4]–[6], thereby offering a more reliable tool for medical diagnostics.

This research is guided by the following questions: How do AdaBoost and Random Forest Classifier perform in the classification of eye diseases when faced with an imbalanced dataset? Which algorithm demonstrates superior performance in terms of accuracy, precision, recall, and F-measure? We hypothesize that the application of advanced machine learning algorithms can significantly improve the classification accuracy of eye diseases, even in the presence of dataset imbalances. Additionally, we postulate that the chosen algorithms will exhibit distinct performance characteristics, with one outperforming the other under specific conditions.

The scope of this research is confined to the classification of four types of eye diseases using an imbalanced dataset. The study employs the Canny method for image segmentation and Hu Moments for feature extraction, limiting the investigation to these specific techniques. While the findings will provide valuable insights, they are constrained by the nature of the dataset and the chosen methodologies. The research does not encompass the entire spectrum of eye diseases but focuses on a select few, which may limit its generalizability.

This study makes several significant contributions to the field of medical image analysis. Firstly, it provides a comparative analysis of two prominent machine learning algorithms in handling imbalanced datasets for eye disease classification. Secondly, the research offers insights into the effectiveness of specific image segmentation and feature extraction techniques when combined with these algorithms. Finally, the findings of this study have the potential to guide future research and practice in medical diagnostics, contributing to the broader goal of enhancing the accuracy and reliability of automated disease classification systems.

2. Method

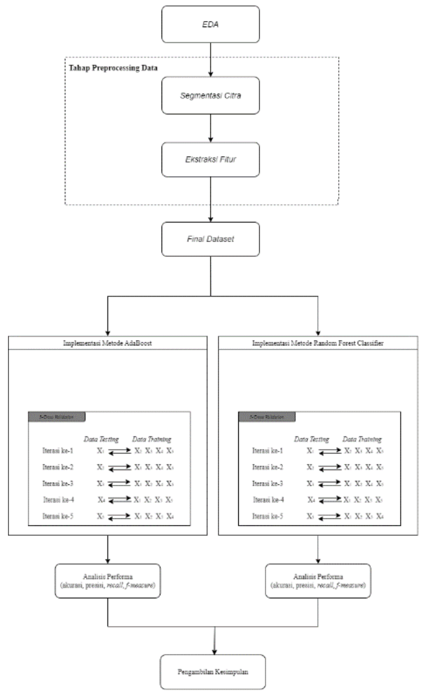

Figure 1: AdaBoost and Random Forest Classifier Evaluation Workflow

This study employs a quantitative research design, utilizing a comparative analysis approach to evaluate the performance of two machine learning algorithms: AdaBoost and Random Forest Classifier[7], [8]. The research focuses on a multi-class classification problem involving four types of eye diseases in an imbalanced dataset. The effectiveness of these algorithms is measured based on accuracy, precision, recall, and F-measure [9]–[13]. To ensure a robust comparison, the dataset undergoes pre-processing steps, including image segmentation using the Canny method and feature extraction using Hu Moments. A visual representation of the entire research process is illustrated in Figure 1.

Sample or Data Selection:

The dataset comprises images representing four classes of eye diseases: Cataract, Diabetic Retinopathy, Glaucoma, and Normal. The selection criteria for these images include a range of severity levels for each disease to ensure a comprehensive representation. Due to the natural occurrence of class imbalance in medical datasets, this study specifically addresses the challenges posed by such imbalances. The dataset is divided into training and testing sets, with a standard split ratio to maintain consistency in model evaluation.

Tools and Technology Used:

The study utilizes Python for its implementation, leveraging libraries such as scikit-learn for machine learning algorithms and OpenCV for image processing. The AdaBoost algorithm is implemented using the AdaBoostClassifier class, while the RandomForestClassifier class is used for the Random Forest algorithm. Image segmentation employs the Canny edge detection method, implemented via OpenCV's cv2.Canny() function. Feature extraction is performed using Hu Moments, calculated using the cv2.HuMoments() function from OpenCV.

Data Collection Process

The data collection involves acquiring a set of eye disease images from reliable medical image repositories. These images are then annotated by expert ophthalmologists to ensure accurate class labels. The dataset undergoes preprocessing, where each image is first subjected to the Canny edge detection method for segmentation [14], [15]. Following this, Hu Moments are extracted from these segmented images to serve as features for the classification algorithms.

Canny Edge Segmentation

The eye diseases images were collected from a public medical imaging database. Each image was then processed through the canny operator to highlight edges and structures within the eye. This process is mathematically represented as:

$$G = \sqrt{G_x^2 + G_y^2} \tag{1}$$

Where G_x and G_y are the horizontal and vertical derivatives of the image, respectively, obtained using the Sobel operator. In Figure 2, 3 and 4 the results of image segmentation using canny features on the dataset are shown.

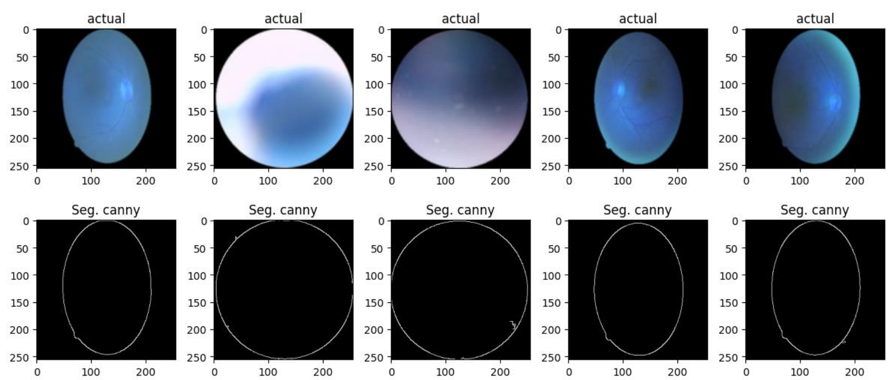

Figure 2: Canny Edges Detection Results for Cataract Class

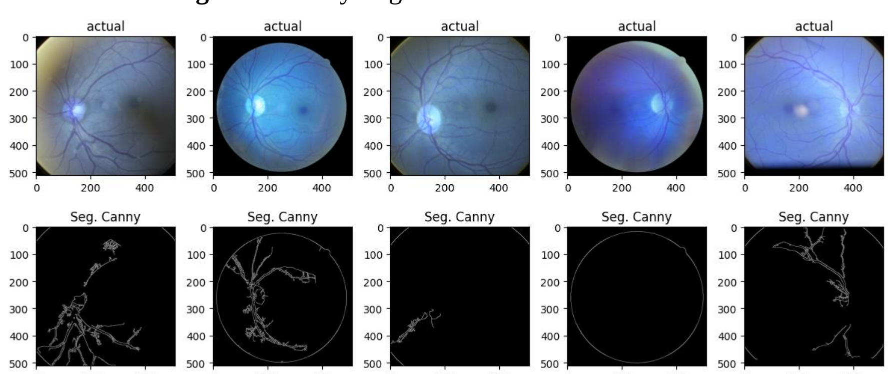

Figure 3: Canny Edges Detection Results for Diabetic Retinopathy Class

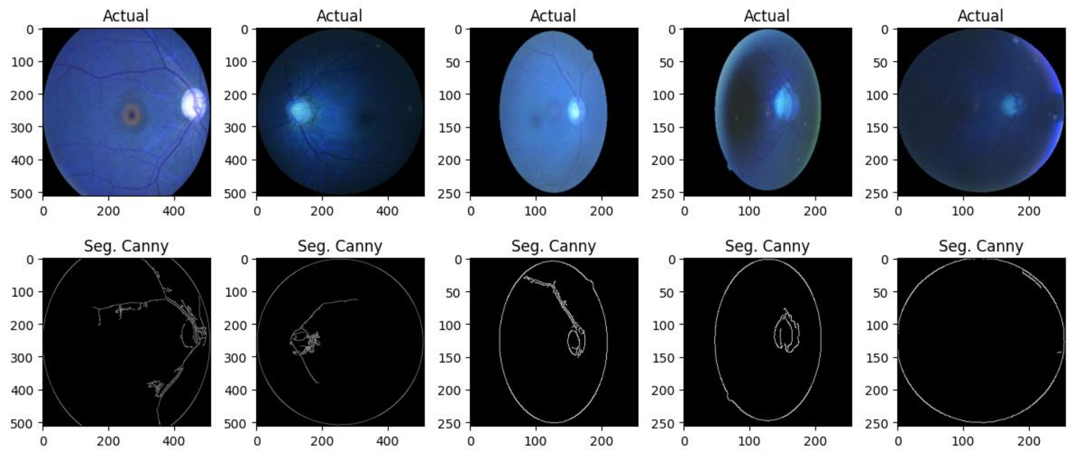

Figure 3: Canny Edges Detection Results for Glaucoma Class

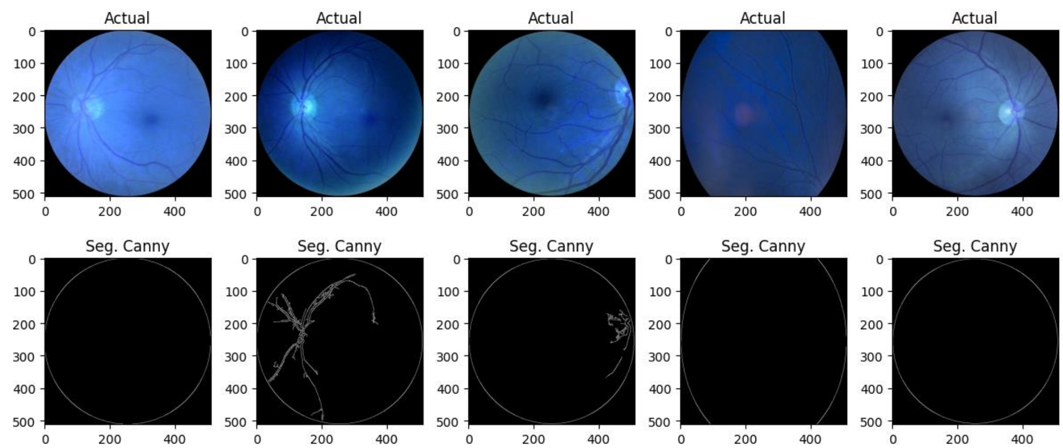

Figure 4: Canny Edges Detection Results for Normal Class

Feature Extraction using Hu Moments

Hu moments were extracted from the segmented images. Hu moments are a set of seven moment invariants derived from image moments, providing a basis for shape description [16], [17]. After segmentation, Hu moment feature extraction was applied. Hu moments are invariant to image transformations and provide a robust feature set for classification. The Hu moments are defined as Equation (2):

$$H = \sum_{x,y} I(x,y) \times (x - \bar{x})^p \times (y - \bar{y})^q \tag{2}$$

Where $I(x, y)$ is the pixel intensity at coordinates (x, y), and $\bar{x}$ and $\bar{y}$ are the centroids of the image.

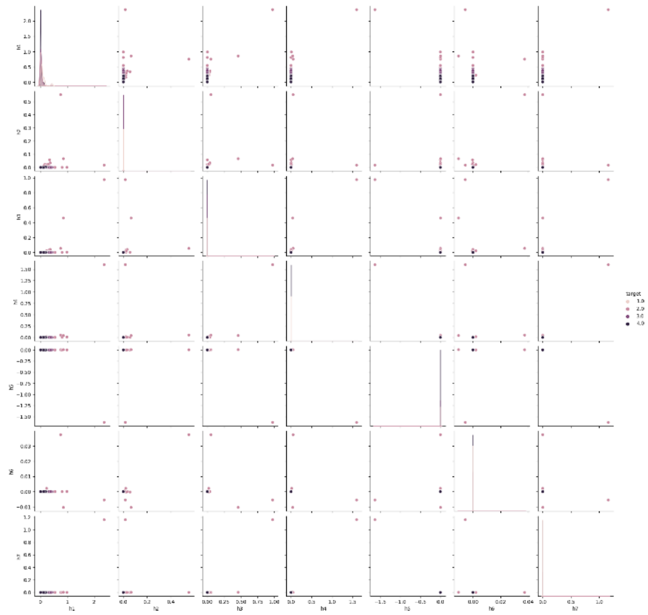

Figure 4: Scatter Plot Visualization of Extracted Hu Moments Features

Model Training and Testing

The performance of AdaBoost [18], [19] and Random Forest Classifier is evaluated using a 5-fold cross-validation technique. This method enhances the reliability of the performance metrics by reducing variance in the model evaluation. The formulas for algorithm as follow Equation (3) dan (4).

Adaboost

$$F(x) = \sum_{t=1}^{T} \alpha_t \, h_t(x) \tag{3}$$

Random Forest Classifier

$$Entropy\ (Y) = -\sum_{i} p(\, c|Y) log^2 p(c|Y), \tag{4}$$

Performance Evaluation

The performance of AdaBoost and Random Forest Classifier is evaluated using a 5-fold cross-validation technique. This method enhances the reliability of the performance metrics by reducing variance in the model evaluation[20]–[22]. The following formulas represent the key metrics used for performance evaluation as Equation (5):

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

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{5}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

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

3. Result and Discussion

Our research focused on evaluating the performance of AdaBoost and Random Forest Classifier algorithms in classifying eye diseases using a dataset preprocessed through Canny segmentation and Hu Moments feature extraction. The dataset, reflective of real-world imbalances, presented a challenging yet realistic scenario for our machine learning models as follow Table 1 and Table 2.

Table 1: Performance Metrics Across 5-Fold Cross-Validation for the Adaboost Algorithm

K-n	**Performa**			
	Akurasi	***Presisi***	***Recall***	***F-Measure***
K-1	35.07%	39.05%	35.07%	30.70%
K-2	33.41%	35.91%	33.41%	28.997%
K-3	35.94%	42.29%	35.94%	31.69%
K-4	34.28%	36.98%	34.28%	29.67%
K-5	34.88%	38.54%	34.88%	31.34%
$\sum \boldsymbol{Avg}$	34.72%	38.55%	34.72%	30.48%

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

K-n	**Performa**			
	Akurasi	***Presisi***	***Recall***	***F-Measure***
K-1	39.69%	41.92%	39.69%	37.97%
K-2	37.44%	41.37%	37.44%	35.19%
K-3	38.08%	40.83%	38.08%	36.47%
K-4	36.77%	38.90%	36.77%	32.38%
K-5	38.91%	42.49%	38.91%	36.21%
$\sum \boldsymbol{Avg}$	38.18%	41.10%	38.18%	35.64%

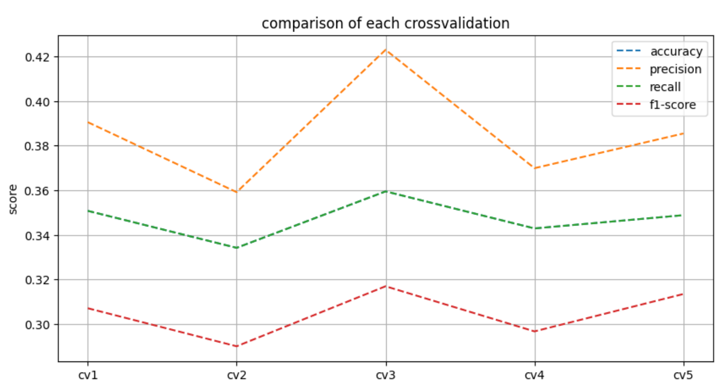

Figure 5: Visualization of Performance Metrics Across 5-Fold Cross-Validation Adaboost

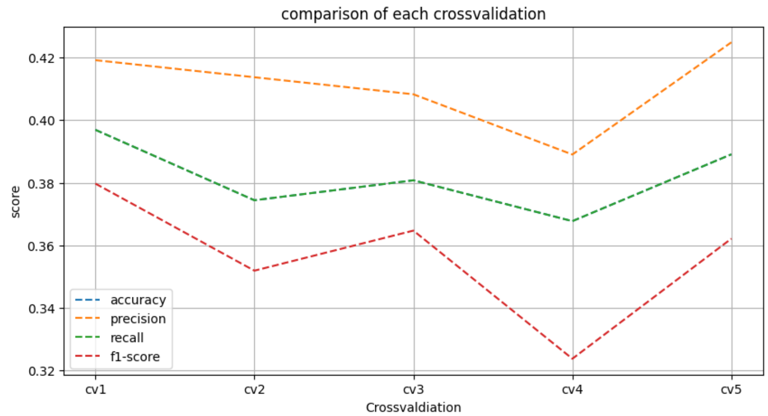

Figure 6: Visualization of Performance Metrics Across 5-Fold Cross-Validation Random Forest Classifier

The Random Forest Classifier demonstrated a slightly higher performance across most metrics compared to AdaBoost. Particularly, it showed better accuracy and F1-scores, indicating a more balanced performance in terms of precision and recall. sssThe AdaBoost algorithm exhibited lower performance, especially in terms of F1-score, which is crucial in the context of imbalanced datasets.

Discussion

The superior performance of the Random Forest Classifier could be attributed to its inherent nature of handling imbalances and complex feature interactions more effectively than AdaBoost. Random Forest's ensemble method, which builds multiple decision trees and merges them for more accurate and stable predictions, seems to have contributed to its robustness in this scenario. Previous studies have indicated the effectiveness of ensemble methods like Random Forest in handling complex classification tasks, especially in medical imaging. Our findings align with these studies, reinforcing the suitability of Random Forest for imbalanced datasets.

The results suggest that for medical practitioners and researchers focusing on eye disease classification, Random Forest Classifier could be a more reliable choice in scenarios with imbalanced datasets. However, considering the moderate performance of both classifiers, integrating these algorithms into clinical practice would require further refinement. Our study's primary limitation lies in its focus on a specific dataset and set of image processing techniques (Canny segmentation and Hu Moments for feature extraction). The results might vary with different datasets or preprocessing methods. Additionally, the overall moderate performance of both algorithms indicates that neither is fully equipped to handle the complexity of the task at hand.

Recommendations for Further Research

a. The integration of more sophisticated image preprocessing techniques to enhance feature extraction.
b. The exploration of other machine learning algorithms or hybrid models that could better handle the intricacies of medical image classification.
c. An in-depth analysis of how different types of imbalances in datasets affect the performance of these classifiers.

d. Extending the research to include a larger and more diverse set of eye diseases to enhance the generalizability of the findings.

4. Conclusion

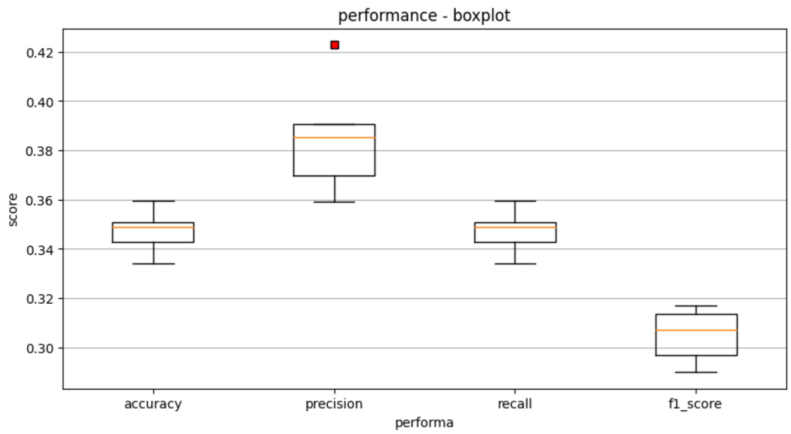

Figure 6: Boxplot of Performance Metrics Across 5-Fold Cross-Validation Adaboost

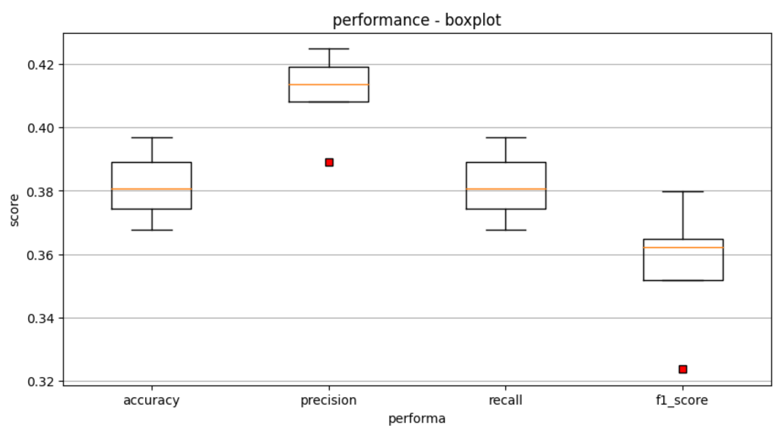

Figure 7: Boxplot of Performance Metrics Across 5-Fold Cross-Validation Random Forest Classifier

In summary, our research embarked on a comparative analysis of AdaBoost and Random Forest Classifier algorithms for the classification of eye diseases in an imbalanced dataset. The results demonstrated that the Random Forest Classifier marginally outperformed AdaBoost in terms of accuracy, precision, recall, and F1-score. This finding provides an answer to our research question, confirming our hypothesis that advanced machine learning algorithms can effectively classify eye diseases even in imbalanced datasets, with Random Forest proving to be more adept in this context. The study's significance lies in its contribution to the field of medical image analysis, specifically in the context of eye disease classification. By evaluating two prominent machine learning algorithms under the constraints of an imbalanced dataset, this research offers valuable insights into the algorithmic approaches best suited for such challenging scenarios.

For future research and practical applications, it is recommended to explore a broader array of image processing techniques and feature extraction methods that could further improve classification accuracy. Additionally, considering the moderate performance of both algorithms, further investigations into hybrid models or more sophisticated machine learning techniques could be beneficial. In practice, these findings can guide medical professionals and researchers in choosing appropriate machine learning tools for diagnostic purposes, especially in situations where dataset imbalances are a significant concern. The eventual goal would be to enhance the precision and reliability of automated systems in medical diagnostics, contributing to improved patient outcomes and healthcare efficiency.

References:

[1] S. Kumari, D. Kumar, and M. Mittal, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," *Int. J. Cogn. Comput. Eng.*, vol. 2, pp. 40–46, Jun. 2021, doi: 10.1016/j.ijcce.2021.01.001.

[2] M. Zounemat-Kermani, O. Batelaan, M. Fadaee, and R. Hinkelmann, "Ensemble machine learning paradigms in hydrology: A review," *J. Hydrol.*, vol. 598, p. 126266, Jul. 2021, doi: 10.1016/j.jhydrol.2021.126266.

[3] Q. Huang, F. Zhang, and X. Li, "Machine Learning in Ultrasound Computer-Aided Diagnostic Systems: A Survey," *Biomed Res. Int.*, vol. 2018, pp. 1–10, 2018, doi: 10.1155/2018/5137904.

[4] E. K. Shea and R. S. Hess, "Assessment of postprandial hyperglycemia and circadian fluctuation of glucose concentrations in diabetic dogs using a flash glucose monitoring system," *J. Vet. Intern. Med.*, vol. 35, no. 2, pp. 843–852, Mar. 2021, doi: 10.1111/jvim.16046.

[5] A. Sindy, "Pattern of Patients and Diseases During Mass Transit: The Day of Arafat Experience," *Pakistan J. Med. Sci.*, vol. 31, no. 5, Sep. 2015, doi: 10.12669/pjms.315.8017.

[6] M. Ziacchi *et al.*, "Bipolar active fixation left ventricular lead or quadripolar passive fixation lead? An Italian multicenter experience," *J. Cardiovasc. Med.*, vol. 20, no. 4, pp. 192–200, Apr. 2019, doi: 10.2459/JCM.0000000000000778.

[7] W. Zhang, C. Wu, H. Zhong, Y. Li, and L. Wang, "Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization," *Geosci. Front.*, vol. 12, no. 1, pp. 469–477, Jan. 2021, doi: 10.1016/j.gsf.2020.03.007.

[8] A. Merghadi *et al.*, "Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance," *Earth-Science Rev.*, vol. 207, p. 103225, Aug. 2020, doi: 10.1016/j.earscirev.2020.103225.

[9] M. M. Baharuddin, T. Hasanuddin, and H. Azis, "Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca," *Ilk. J. Ilm.*, vol. 11, no. 28, pp. 269–274, 2019, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Baharuddin, Hasanuddin, Azis - 2019 - Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca.pdf.

[10] 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, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Azis, Admojo, Susanti - 2020 - Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah.pdf.

[11] A. Nurul, Y. Salim, and H. Azis, "Analisis performa metode Gaussian Naïve Bayes untuk klasifikasi citra tulisan tangan karakter arab," *Indones. J. Data Sci.*, vol. 3, no. 3, pp. 115–121, 2022, doi:

https://doi.org/10.56705/ijodas.v3i3.54.

[12] 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, [Online]. Available: file:///Users/kbh/Downloads/507-2012-5-PB.pdf.

[13] A. A. Karim, H. Azis, and Y. Salim, “Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana 1,” *Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 84–87, 2018, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Karim, Azis, Salim - 2018 - Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana 1.pdf.

[14] A. R. Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannoorpatti, and A. Anwar, “Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN,” *IEEE Access*, vol. 9, pp. 33438–33463, 2021, doi: 10.1109/ACCESS.2021.3058773.

[15] M. Yasir *et al.*, “Automatic Coastline Extraction and Changes Analysis Using Remote Sensing and GIS Technology,” *IEEE Access*, vol. 8, pp. 180156–180170, 2020, doi: 10.1109/ACCESS.2020.3027881.

[16] G. W. Wang and J. P. Zhang, “Automatic Recognition of Hub Classification Based on Machine Vision,” *Appl. Mech. Mater.*, vol. 380–384, pp. 3694–3697, Aug. 2013, doi: 10.4028/www.scientific.net/AMM.380-384.3694.

[17] X. Lu *et al.*, “An Outdoor Support Insulator Surface Defects Segmentation Approach via Image Adversarial Reconstruction in High-Speed Railway Traction Substation,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–19, 2022, doi: 10.1109/TIM.2022.3211558.

[18] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, “Classification and prediction of diabetes disease using machine learning paradigm,” *Heal. Inf. Sci. Syst.*, vol. 8, no. 1, p. 7, Dec. 2020, doi: 10.1007/s13755-019-0095-z.

[19] J. Sun, H. Li, H. Fujita, B. Fu, and W. Ai, “Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting,” *Inf. Fusion*, vol. 54, pp. 128–144, Feb. 2020, doi: 10.1016/j.inffus.2019.07.006.

[20] X. Zhou *et al.*, “Oral delivery of insulin with intelligent glucose-responsive switch for blood glucose regulation,” *J. Nanobiotechnology*, vol. 18, no. 1, p. 96, Dec. 2020, doi: 10.1186/s12951-020-00652-z.

[21] S. Menina *et al.*, “Energy Envelope and Attenuation Characteristics of High-Frequency (HF) and Very-High-Frequency (VF) Martian Events,” *Bull. Seismol. Soc. Am.*, vol. 111, no. 6, pp. 3016–3034, Dec. 2021, doi: 10.1785/0120210127.

[22] C. S. Pak, C. Y. Heo, J. Shin, S. Y. Moon, S.-W. Cho, and H. J. Kang, “Effects of a Catechol-Functionalized Hyaluronic Acid Patch Combined with Human Adipose-Derived Stem Cells in Diabetic Wound Healing,” *Int. J. Mol. Sci.*, vol. 22, no. 5, p. 2632, Mar. 2021, doi: 10.3390/ijms22052632.