International Journal of Artificial Intelligence in Medical Issues

Volume 2 Issue 2
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v2i2.176

*Research Article*

# Classification of Pseudopapilledema and Papilledema Using Decision Tree and Hu Moments

**Hayatou Oumarou** [1,*]
[1] The University of Maroua, Cameroon, hayaty55@yahoo.fr
Correspondence should be addressed to Hayatou Oumarou; hayaty55@yahoo.fr

Received 25 September 2024; Revised 03 October 2024; Accepted 20 November 2024; Published 30 November 2024



**Abstract:**
Pseudopapilledema, characterized by an anomalous elevation of the optic disc without retinal nerve fiber layer edema, often mimics the presentation of true papilledema caused by increased intracranial pressure. Accurate differentiation between these conditions is critical to avoid unnecessary invasive procedures. This study employs a Decision Tree classifier to classify optic disc images into three categories: normal, papilledema, and pseudopapilledema. The dataset, obtained from Kaggle, consists of imbalanced images segmented using the Canny edge detection method and features extracted using Hu Moments. The dataset was divided into 80% training and 20% testing sets. Performance was evaluated using 5-fold cross-validation, yielding an average accuracy of 53.61%, precision of 55.20%, recall of 54.12%, and F1-score of 55.17%. The study provides a comprehensive analysis of the classifier's performance, including visualizations such as segmentation results, scatter plots of Hu Moments, and confusion matrices. The results indicate that while the Decision Tree classifier demonstrates moderate effectiveness, there is significant room for improvement. The research highlights the potential of machine learning models in medical diagnostics but also underscores the need for more robust algorithms and diverse datasets. Future work should focus on incorporating more complex models and expanding the dataset to enhance diagnostic accuracy. These findings contribute to the field of medical image analysis and propose a non-invasive diagnostic tool that, when integrated with clinical expertise, could improve patient outcomes and reduce unnecessary procedures.

**Keywords:** Pseudopapilledema, Papilledema, Decision Tree, Hu Moments, Medical Diagnostics.

**Dataset link:** https://www.kaggle.com/datasets/shashwatwork/identification-of-pseudopapilledema

## 1. Introduction

Pseudopapilledema is characterized by an anomalous elevation of one or both optic discs without the presence of edema in the retinal nerve fiber layer. This condition often presents similarly to papilledema, which is the swelling of the optic disc due to increased intracranial pressure. Distinguishing between pseudopapilledema and true papilledema is crucial, as the latter can be an early sign of serious conditions such as optic disc diseases, which may lead to vision loss, neurological impairment, or even death. Accurate differentiation is essential to avoid unnecessary invasive procedures and to ensure appropriate treatment. The advent of machine learning and image processing technologies offers new opportunities to improve diagnostic accuracy in this area [1], [2].

The primary problem addressed in this research is the misdiagnosis of pseudopapilledema as papilledema. Misdiagnosis can lead to patients undergoing unnecessary lumbar punctures, MRI scans, and extensive laboratory tests, which are not only invasive but also costly and time-consuming. Accurate, non-invasive diagnostic tools are needed to differentiate these conditions effectively. The availability of a dataset on Kaggle, comprising images of normal optic discs, papilledema, and pseudopapilledema, provides an opportunity to develop and test such tools using

machine learning techniques. The objective of this study is to develop a Decision Tree classifier capable of accurately distinguishing between normal, papilledema, and pseudopapilledema optic disc images. The classifier will utilize image segmentation via the Canny method and feature extraction using Hu Moments [3], [4]. The dataset will be pre-processed and split into training and testing sets, with performance evaluated based on accuracy, precision, recall, and F1-measure [5], [6]. By achieving these objectives, the study aims to contribute to the field of medical diagnostics with a reliable and non-invasive tool for differentiating optic disc conditions.

This research aims to answer several key questions: How accurately can a Decision Tree classifier distinguish between normal, papilledema, and pseudopapilledema based on the given dataset? What are the specific performance metrics (accuracy, precision, recall, and F1-measure) of the classifier? Additionally, the study seeks to explore the practical implications of using such a classifier in clinical settings. By addressing these questions, the research contributes to the broader field of medical image analysis and diagnostic machine learning. The scope of this research is limited to the dataset obtained from Kaggle, which includes images collected from the internet and categorized into three classes. While the dataset provides a useful starting point, it may not encompass all possible variations seen in clinical practice. Additionally, the study focuses on a single classification method, the Decision Tree, although future research could explore other machine learning algorithms for potentially better performance. Despite these limitations, the study offers a significant step towards non-invasive diagnostic tools for optic disc conditions.

The contributions of this research are manifold. First, it demonstrates the feasibility of using machine learning techniques, specifically Decision Trees, for medical image classification. Second, it provides a detailed methodology for pre-processing and feature extraction that can be applied to similar datasets. Third, the study's findings have practical implications for reducing unnecessary invasive procedures in clinical diagnostics. Finally, the research opens avenues for further studies to refine and expand the classifier, potentially incorporating additional data and exploring other machine learning models to enhance diagnostic accuracy.

## 2. Method

The research employs an experimental design approach, utilizing machine learning techniques to classify and differentiate between normal, papilledema, and pseudopapilledema optic disc images. The primary method involves image segmentation using the Canny edge detection method and feature extraction using Hu Moments [4]. The dataset is split into training and testing sets, followed by the application of a Decision Tree classifier [7]. The performance of the classifier is evaluated using metrics such as accuracy, precision, recall, and F1-measure [8], [9]. This structured approach ensures a comprehensive analysis of the classifier's effectiveness in differentiating the specified conditions. A visual representation of the entire research process is illustrated in **Figure 1**.

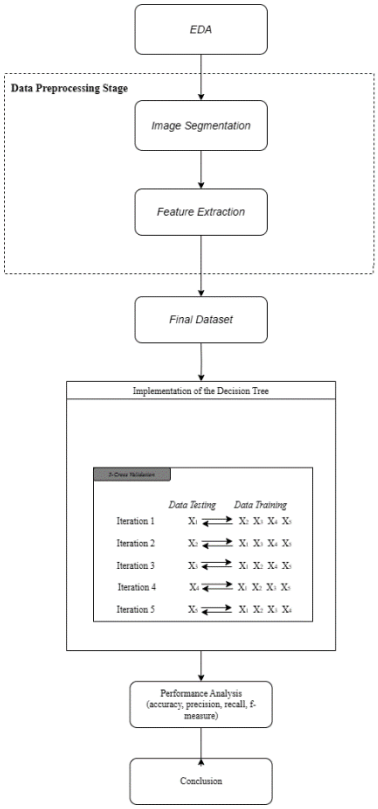


**Figure 1:** Decision Tree Evaluation Workflow

**Sample or Data Selection:**

The dataset used in this study is obtained from Kaggle and comprises images of optic discs categorized into three classes: normal, papilledema, and pseudopapilledema. The dataset is inherently imbalanced, reflecting the real-world prevalence of these conditions. To ensure robust model training and evaluation, the dataset is split into 80% training and 20% testing sets [10]. The splitting process is conducted using stratified sampling to maintain the class distribution in both subsets.

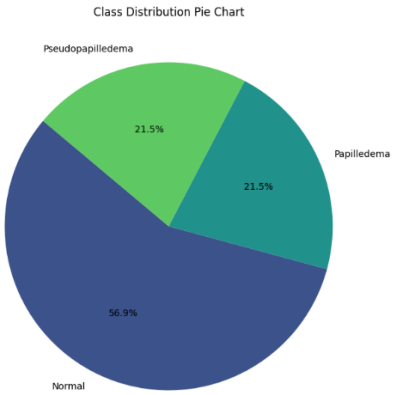


**Figure 2:** Class Distribution

**Data Collection Process**

The images in the dataset are collected from various internet sources and categorized based on medical diagnosis into the three specified classes. Each image undergoes pre-processing steps including resizing, normalization, and conversion to grayscale to standardize the input data for subsequent analysis. Image segmentation is performed using the Canny edge detection method to highlight the structural features of the optic disc. Following segmentation, feature extraction is carried out using Hu Moments, which are invariant to image transformations such as rotation, scale, and translation.

**Data Analysis Methods**

The data analysis process involves several steps to ensure accurate classification and evaluation of the model:

a. Pre-processing:

- Resizing images to a uniform size.
- Normalizing pixel values to a range of [0, 1].
- Converting images to grayscale.
- Segmenting images using the Canny edge detection method [11]–[13]:

$$Edges = Canny(Image, threshold1, threshold2) \quad (1)$$

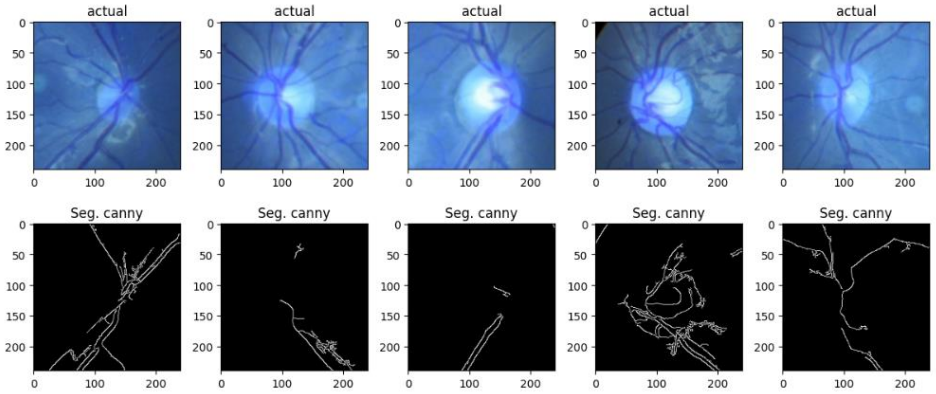


**Figure 3:** Class Normal

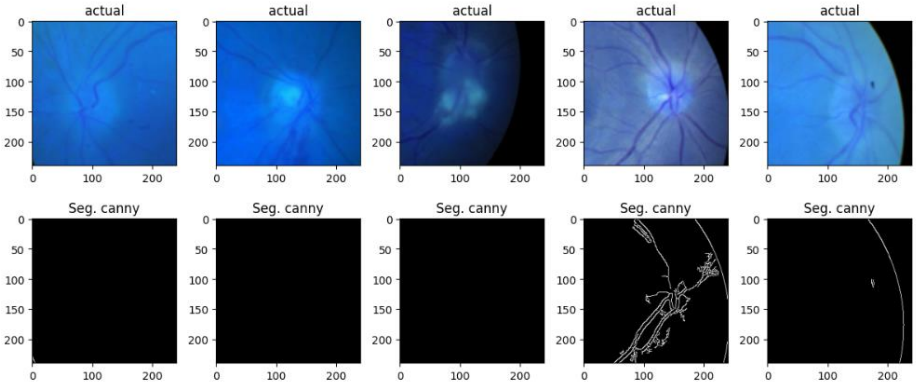


**Figure 4:** Class Papilledema

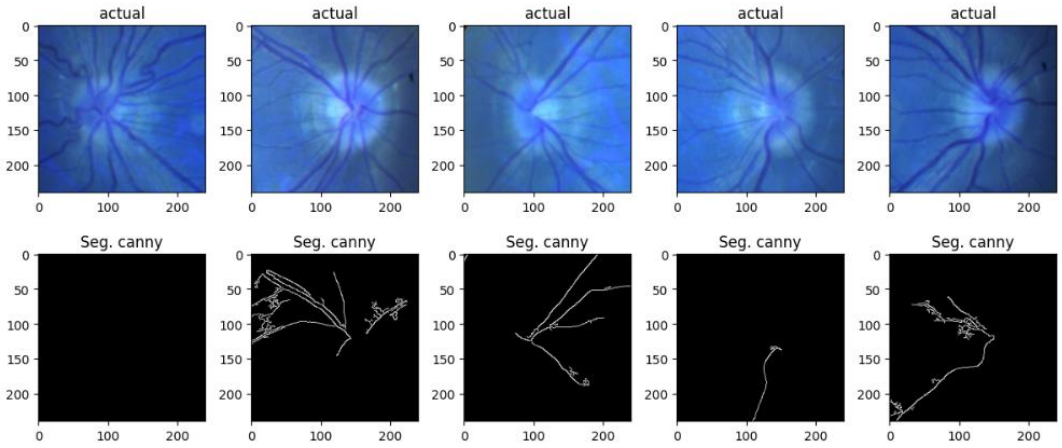


**Figure 5:** Class Pseudopapilledema

**Figure 3** to **5** shows the output of the Canny edge detection on sample images from each class, highlighting the edges of the optic discs

b. Feature Extraction: Calculating Hu Moments for each segmented image [14]–[16]. Hu Moments are calculated using the following invariant moments [17]–[19]:

$$\begin{aligned} H_1 &= \mu_{20} + \mu_{02} \\ H_2 &= (\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2 \\ &\vdots \\ H_7 &= \mu_{30}\mu_{12} - \mu_{21}\mu_{03} - 3\mu_{12}^2\mu_{03} + 3\mu_{21}^2\mu_{12} \end{aligned} \tag{3}$$

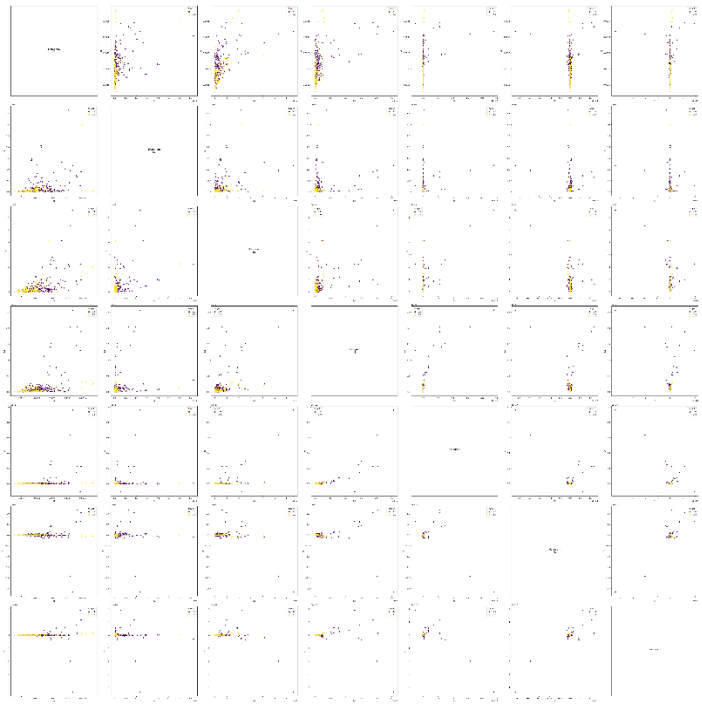

**Figure 6:** Scatter Plots for All Combinations of Hu Moments

**Figure 6** plots display the pairwise relationships between the extracted Hu Moments, providing insights into their distribution and potential separability of the classes.

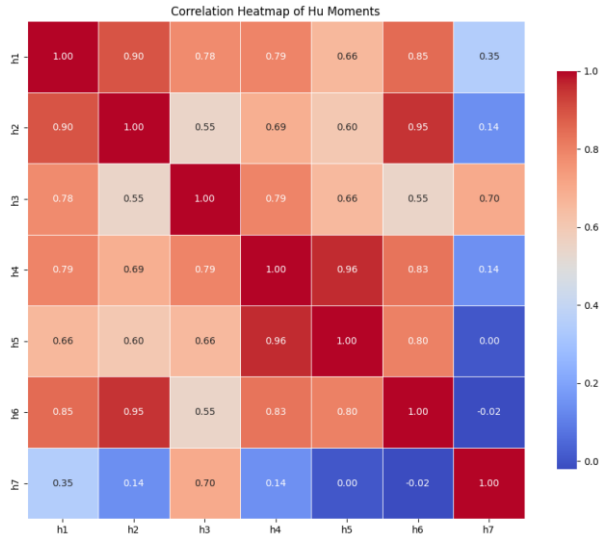


**Figure 7:** Correlation Heatmap of Hu Moments

**Figure 7** heatmap displays the correlation coefficients between the Hu Moments, indicating the degree of linear relationship between the features.

c. Scaling: Standardizing the features to have mean 0 and variance 1.

$$scaled_{Ferature} = \frac{feature - \mu}{\sigma} \tag{3}$$

Where $\mu$ is the mean and $\sigma$ is the standard deviation of the feature.

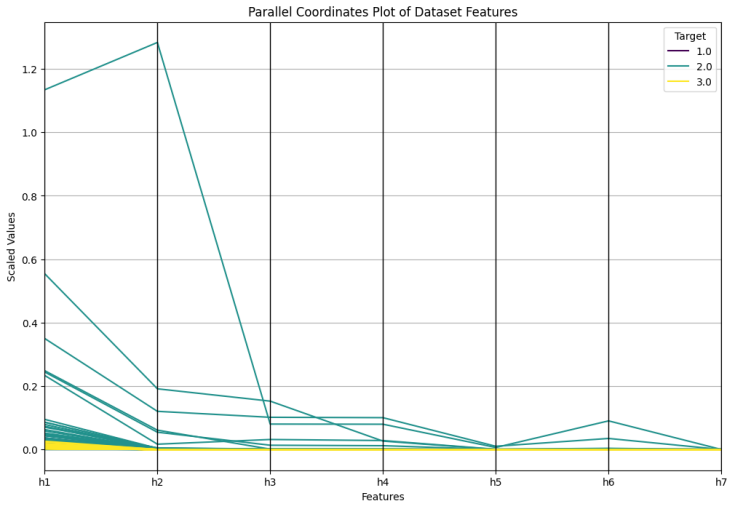


**Figure 8:** Parallel Coordinates Plot of Dataset Features

**Figure 8** plot shows each feature across multiple dimensions, helping to visualize how different features vary together across the classes.

d. Model Implementation:

- Splitting Data:

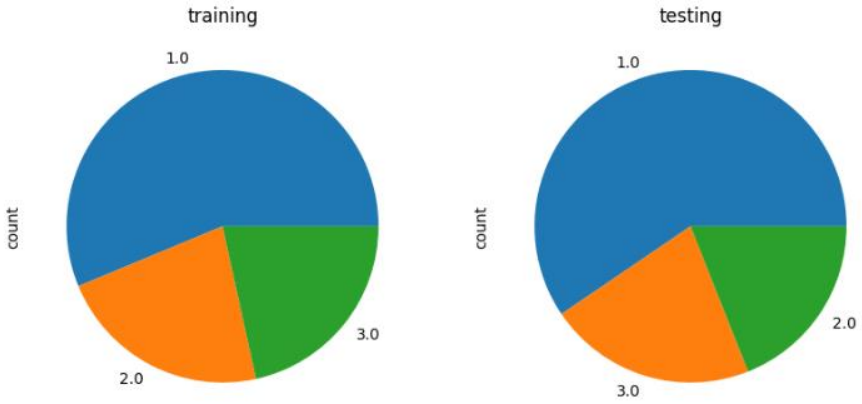


**Figure 2:** Splitting Data Training (80%), Testing (20%)

- Training the model on the training set and evaluating on the testing set.

  The Decision Tree algorithm [20] works by recursively splitting the dataset based on a feature that results in the best possible classification [21], [22]. The splitting criterion is often based on metrics such as Gini impurity or Information Gain (entropy) [23]–[25].

  **Gini Impurity:**

$$Gini(D) = 1 - \sum_{i=1}^{m} p_i^2 \tag{4}$$

  Where $p_i$ is the probability of a randomly chosen element being classified correctly in class $i$.

  **Information Gain (Entropy):**

$$Gini(D, A) = Entropy(D) - \sum_{v \in Values(A)} \frac{|D_v|}{|D|} Entropy(D_v) \tag{5}$$

  where

$$Entropy(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \tag{6}$$

  and $D_v$ is the subset of $D$ for which feature $A$ has value $v$.

e. Performance Evaluation: Calculating accuracy, precision, recall, and F1-measure [10], [26]–[28].

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

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{7}$$

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

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

## 3. Result and Discussion

The dataset was pre-processed to ensure uniformity, involving resizing of images, normalization of pixel values to the range [0, 1], and conversion to grayscale. Following this, image segmentation was performed using the Canny edge detection method, and feature extraction was conducted using Hu Moments. The dataset was split into training (80%) and testing (20%) sets, and a Decision Tree classifier was implemented to categorize the optic disc images into normal, papilledema, and pseudopapilledema.

The performance of the Decision Tree classifier was evaluated using 5-fold cross-validation. The results are summarized in the following **Table 1**.

**Table 1:** Performance Metrics Across 5-Fold Cross-Validation for the Decision Tree Algorithm

| K-n | Metrics | | | |
|---|---|---|---|---|
| | ***Accuracy*** | ***Precision*** | ***Recall*** | ***F-Measure*** |
| K-1 | 51.82% | 53.28% | 51.09% | 51.83% |
| K-2 | 51.82% | 51.9% | 50.73% | 52.58% |
| K-3 | 54.01% | 56.04% | 56.2% | 56.27% |
| K-4 | 55.47% | 58.23% | 59.85% | 58.73% |
| K-5 | 54.95% | 56.57% | 52.75% | 56.45% |
| $\sum$ ***Avg*** | 53.61% | 55.2% | 54.12% | 55.17% |

To better understand the performance of the model, we include graphical visualizations such as the performance metrics over the folds and the confusion matrix for the classifier's predictions.

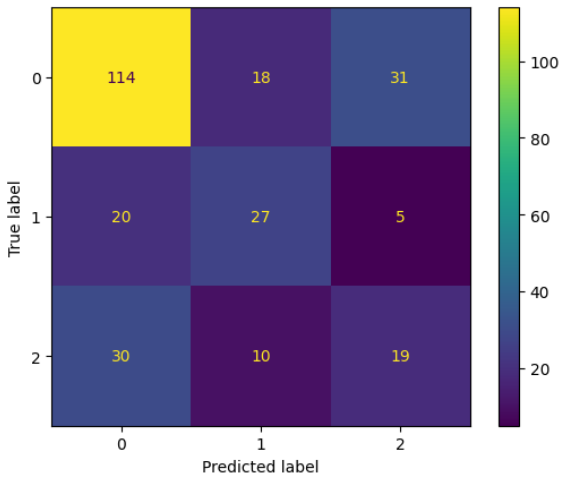


**Figure 9:** Confusion Matrix of the Decision Tree

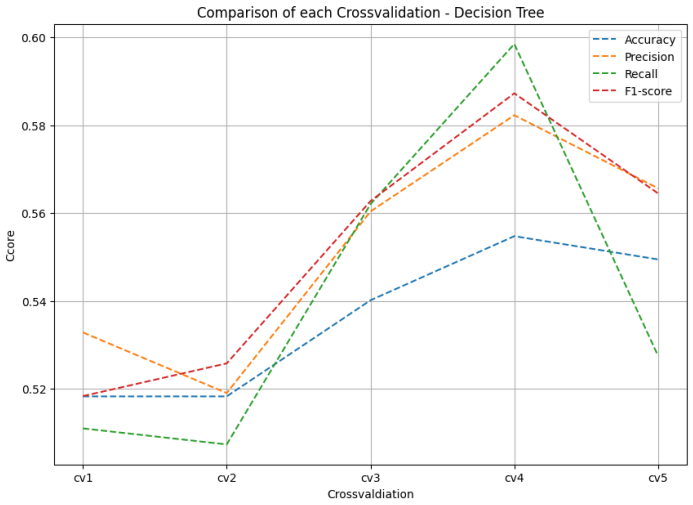


**Figure 10:** Performance Comparison of Each Cross-validation Fold

The analysis of the Decision Tree classifier reveals that the model's performance is consistent across different folds, with slight variations in accuracy, precision, recall, and F1-score. The highest accuracy achieved was 55.47%, and the highest precision was 58.23%, indicating the classifier's moderate success in correctly identifying the conditions. However, the recall rate of 59.85% in one-fold suggests that while the classifier can identify the conditions well, it may not consistently perform at this level.

**Discussion**

The results indicate that the Decision Tree classifier, while moderately effective, has room for improvement. The classifier's mean accuracy of 53.61% suggests that more than half of the predictions are correct, but there is a significant margin for enhancement. The precision and recall rates further support the need for refinement, with mean values of 55.20% and 54.12%, respectively, highlighting the trade-offs between identifying true positives and false positives. Comparing these findings with previous research, it is evident that while machine learning models like Decision Trees can provide valuable diagnostic tools, they may not always offer the highest accuracy without additional tuning or the use of more complex algorithms. Studies have shown that ensemble methods or deep learning approaches might yield better performance due to their ability to capture more intricate patterns in the data.

The practical implications of these results underscore the potential for using Decision Trees in clinical settings as a non-invasive diagnostic aid. However, the current performance levels suggest that relying solely on this model could lead to misdiagnoses. Therefore, it is crucial to consider these models as part of a broader diagnostic toolkit, complemented by clinical expertise and additional tests. The limitations of this research include the imbalanced dataset and the potential variability in the image quality sourced from the internet. Future research should focus on expanding the dataset to include more diverse and clinically validated images. Additionally, exploring other machine learning techniques such as Random Forests, Support Vector Machines, or Convolutional Neural Networks could provide more robust performance.

## 4. Conclusion

This study demonstrates the application of a Decision Tree classifier to distinguish between normal, papilledema, and pseudopapilledema optic disc images. The results reveal a moderate accuracy of 53.61%, with precision and recall rates of 55.20% and 54.12%, respectively, highlighting the model's potential yet indicating the need for further refinement. The classifier's ability to achieve a mean F1-score of 55.17% across different folds confirms its capability to differentiate the conditions to some extent, but also underscores the necessity for improvement.

The research successfully addresses the primary objective of developing a non-invasive diagnostic tool for differentiating optic disc conditions, contributing valuable insights into the use of machine learning in medical image analysis. Despite its limitations, this study lays the groundwork for future research, recommending the exploration of more complex algorithms and larger, more diverse datasets to enhance diagnostic accuracy. Integrating these models into a comprehensive diagnostic framework alongside clinical expertise could significantly improve patient outcomes and reduce unnecessary invasive procedures.

**References:**


[1] Z. H. Zhou, *Machine Learning*. 2021.

[2] M. R. Sharan, "Classification of Medicinal Leaf by Using Canny Edge Detection and SVM Classifier," *2022 Int. Conf. Futur. Technol. INCOFT 2022*, 2022, doi: 10.1109/INCOFT55651.2022.10094461.

[3] 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.

[4] 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.

[5] D. İzci, "Comparative performance analysis of slime mould algorithm for efficient design of proportional–integral–derivative controller," *Electrica*, vol. 21, no. 1, pp. 151–159, 2021, doi: 10.5152/ELECTRICA.2021.20077.

[6] M. Khushi, "A Comparative Performance Analysis of Data Resampling Methods on Imbalance Medical Data," *IEEE Access*, vol. 9, pp. 109960–109975, 2021, doi: 10.1109/ACCESS.2021.3102399.

[7] R. Sun, "A gradient boosting decision tree based GPS signal reception classification algorithm," *Appl. Soft Comput. J.*, vol. 86, 2020, doi: 10.1016/j.asoc.2019.105942.

[8] H. Azis, D. Widyawati, and ..., "Prediksi potensi donatur menggunakan model Logistic Regression," *Indones. J. ...*, 2023, [Online]. Available: https://jurnal.yoctobrain.org/index.php/ijodas/article/view/64.

[9] 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.

[10] U. Zaky, A. Naswin, S. Sumiyatun, and ..., "Performance Analysis of the Decision Tree Classification Algorithm on the Water Quality and Potability Dataset," *Indones. J. ...*, 2023.

[11] S. K. Jadwaa, "X-Ray Lung Image Classification Using a Canny Edge Detector," *J. Electr. Comput. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/3081584.

[12] N. A. B. Mary, "Classification of banana leaf diseases using enhanced gabor feature descriptor," *Lecture Notes in Networks and Systems*, vol. 145. pp. 229–242, 2021, doi: 10.1007/978-981-15-7345-3_19.

[13] K. U. Devi, "Brain tumour classification using saliency driven nonlinear diffusion and deep learning with convolutional neural networks (CNN)," *Journal of Ambient Intelligence and Humanized Computing*. 2020, doi: 10.1007/s12652-020-02200-x.

[14] A. Abisha, "Feature Extraction from Plant Leaves and Classification of Plant Health Using Machine Learning," *Lecture Notes in Electrical Engineering*, vol. 858. pp. 867–876, 2022, doi: 10.1007/978-981-19-0840-8_67.

[15] Y. Benlachmi, "Fruits Disease Classification using Machine Learning Techniques," *Indones. J. Electr. Eng. Informatics*, vol. 10, no. 4, pp. 917–929, 2022, doi: 10.52549/ijeei.v10i4.3907.

[16] S. Kavitha, "Classification of Indian Monument Architecture Styles Using Bi-Level Hybrid Learning Techniques," *Lecture Notes in Networks and Systems*, vol. 436. pp. 471–488, 2022, doi: 10.1007/978-981-19-1012-8_32.

[17] A. Naswin and A. P. Wibowo, "Performance Analysis of the Decision Tree Classification Algorithm on the Pneumonia Dataset," *... Artif. Intell. Med. ...*, 2023.

[18] 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.

[19] R. A. Azdy, R. F. Syam, E. Faizal, and ..., "Performance Evaluation of Bagging Meta-Estimator in Lung Disease Detection: A Case Study on Imbalanced Dataset," *Int. J. ...*, 2023.

[20] H. Azis and S. R. Jabir, "Chemical Composition and Aroma Profiling: Decision Tree Modeling of Formalin Tofu," *J. Embed. Syst. Secur. ...*, 2023.

[21] S. H. Asman, "Decision tree method for fault causes classification based on rms-dwt analysis in 275 kv transmission lines network," *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11094031.

[22] F. Manzella, "The voice of COVID-19: Breath and cough recording classification with temporal decision trees and random forests," *Artif. Intell. Med.*, vol. 137, 2023, doi: 10.1016/j.artmed.2022.102486.

[23] M. A. Hafeez, "Performance improvement of decision tree: A robust classifier using tabu search algorithm," *Appl. Sci.*, vol. 11, no. 15, 2021, doi: 10.3390/app11156728.

[24] D. Jalal, "Decision Tree and Support Vector Machine for Anomaly Detection in Water Distribution Networks," *2020 International Wireless Communications and Mobile Computing, IWCMC 2020*. pp. 1320–1323, 2020, doi: 10.1109/IWCMC48107.2020.9148431.

[25] R. Hazra, "Machine Learning for Breast Cancer Classification with ANN and Decision Tree," *11th Annual IEEE Information Technology, Electronics and Mobile Communication Conference, IEMCON 2020*. pp. 522–527, 2020, doi: 10.1109/IEMCON51383.2020.9284936.

[26] G. Giri, I. A. Musdar, H. Angriani, and ..., "Enhancing Disease Management in Mango Cultivation: A Machine

Learning Approach to Classifying Leaf Diseases," *Indones. J. ...*, 2023.

[27] A. Faradibah, D. Widyawati, A. U. T. Syahar, and ..., "Comparison Analysis of Random Forest Classifier, Support Vector Machine, and Artificial Neural Network Performance in Multiclass Brain Tumor Classification," *Indones. J. ...*, 2023.

[28] H. A. Siregar, M. Z. Raditya, A. N. Yesa, and ..., "Comparison of Classification Algorithm Performance for Diabetes Prediction Using Orange Data Mining," *Indones. J. ...*, 2023.