International Journal of Artificial Intelligence in Medical Issues

Volume 2 Issue 1
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v2i1.140

Research Article

Analysing Musculoskeletal Bone Fractures Using Decision Trees: A Deep Learning Approach with Canny Segmentation and Hu Moments Feature Extraction

Riska [1,*]
[1] Politeknik Negeri Ujung Pandang, riskaanasirr@gmail.com
Correspondence should be addressed to Riska; riskaanasirr@gmail.com

Received 26 March 2024; Revised 17 April 2024; Accepted 10 May 2024; Published 31 May 2024

Abstract:
This study presents an in-depth analysis of the application of a Decision Tree classifier to detect bone fractures from X-ray images, leveraging the FracAtlas dataset containing 4,083 labelled images. The classifier underwent a rigorous evaluation using 5-fold cross-validation, focusing on metrics such as accuracy, precision, recall, and F1-score to ascertain its performance. Results varied across folds, with an accuracy range of 69.89% to 74.05%, precision between 72.27% and 73.75%, recall from 70.50% to 73.81%, and F1-scores of 71.52% to 73.31%. A graphical depiction of these metrics provided a visual comparison of performance consistency, while the confusion matrix offered a detailed account of the model’s predictive success and shortcomings. The research confirms the hypothesis that integrating Canny edge detection for segmentation and Hu Moments for feature extraction with a Decision Tree approach can facilitate fracture identification, positing the model as a supportive tool for radiologists. The study's findings contribute to the field of medical image analysis, suggesting that machine learning can be a valuable asset in clinical diagnostics. Recommendations for future research include the exploration of more complex algorithms, expansion of the dataset, and refinement of pre-processing techniques, to enhance the model's diagnostic precision further.

Keywords: Bone Fracture Detection, Decision Tree, Machine Learning, Medical Image Analysis, Radiology.

Dataset link: https://www.kaggle.com/datasets/francismon/curated-colon-dataset-for-deep-learning

1. Introduction

In the realm of medical diagnostics, the swift and accurate identification of musculoskeletal injuries, particularly bone fractures, is a pivotal challenge. As the prevalence of such injuries escalates globally, fuelled by both aging populations and increasing participation in high-risk sports, the demand for more efficient diagnostic methodologies has never been more critical. Traditional approaches, relying predominantly on the keen eyes and vast experience of radiologists, face limitations in consistency and objectivity, often leading to variability in diagnostic outcomes. This backdrop underscores the imperative need for advancements in automated diagnostic technologies, which promise to augment the accuracy and reliability of fracture detection, thereby facilitating improved patient care and treatment planning.

The core problem that this research seeks to address revolves around the current limitations inherent in manual fracture diagnosis processes. Despite advancements in imaging technologies, the interpretation of these images remains a highly specialized skill, susceptible to human error and subjectivity. This scenario is further complicated by the sheer volume of cases that medical professionals must evaluate, potentially leading to diagnostic delays and decreased throughput in healthcare facilities. The introduction of machine learning and image processing techniques offers a promising solution to these challenges, aiming to support and enhance the diagnostic capabilities of radiologists by providing them with automated tools that can rapidly and accurately identify fractures in X-ray images.

The primary objective of this study is to investigate the efficacy of a Decision Tree-based machine learning model, augmented by Canny edge detection for image segmentation and Hu Moments [1]–[3] for feature extraction, in classifying bone fractures from X-ray images. By leveraging the FracAtlas dataset, which encompasses a comprehensive collection of annotated X-ray images, this research aims to demonstrate how advanced computational techniques can be harnessed to improve the accuracy and efficiency of bone fracture diagnosis.

This research is guided by the hypothesis that the integration of machine learning algorithms with sophisticated image pre-processing techniques can significantly enhance the precision of fracture detection in musculoskeletal X-rays. The study seeks to answer several critical questions: Can a Decision Tree model [4]–[6], informed by features extracted through Canny segmentation [7], [8] and Hu Moments, outperform traditional diagnostic methods in terms of accuracy, precision, recall, and F1-measure [9]–[11].

While the scope of this research is focused on the classification of bone fractures using a specific dataset and set of algorithms, it is important to acknowledge the limitations that accompany this study. The model's performance is inherently dependent on the quality and diversity of the dataset, which may not fully encapsulate the vast variability of fracture presentations encountered in clinical practice. Additionally, the study assumes a certain level of accuracy in the initial annotations of the FracAtlas dataset, which, if flawed, could impact the validity of the research findings.

Notwithstanding these limitations, the contributions of this research are manifold. By systematically evaluating the performance of a Decision Tree algorithm in conjunction with Canny segmentation and Hu Moments feature extraction, this study not only adds to the burgeoning field of automated medical diagnostics but also provides a framework for future investigations into the application of machine learning in radiology. Furthermore, the insights gleaned from this research have the potential to inform the development of new tools and technologies that could revolutionize the way fractures are diagnosed, ultimately leading to faster, more accurate, and more objective diagnostic processes.

2. Method

Our study employs a quantitative research design, focusing on the application of a Decision Tree classifier [12], [13] to a pre-processed dataset of X-ray images. The pre-processing involves Canny edge detection for segmentation and Hu Moments for feature extraction, aimed at isolating and identifying relevant features for fracture classification. The research design encompasses the application of 5-fold cross-validation [14] to evaluate the model's performance, ensuring a robust assessment across different subsets of the dataset. A visual representation of the entire research process is illustrated in **Figure 1**.

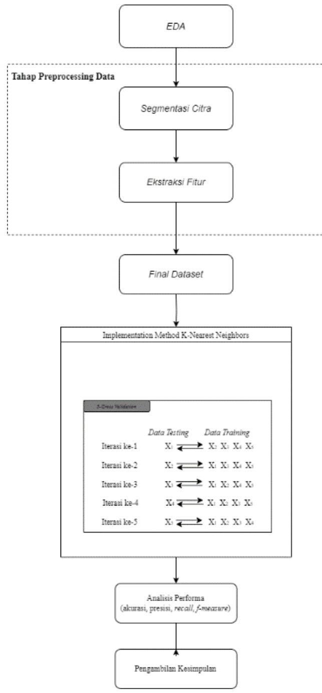

Figure 1: Decision Tree Evaluation Workflow

Sample or Data Selection:

The dataset utilized in this research, FracAtlas, comprises 4,083 X-ray images annotated for the presence or absence of bone fractures. Each image is labelled as either 1 (Fractured) or 2 (Non-Fractured), serving as the basis for classification. The dataset is partitioned into training and test sets in a manner that ensures balanced representation of both classes in each fold of the cross-validation process.

Tools and Technology Used:

The research employs several tools and technologies:

- Python: For implementing the Decision Tree algorithm and preprocessing steps.
- OpenCV: Used for applying Canny edge detection to segment the images.
- Scikit-learn: Utilized for the Decision Tree classifier and cross-validation.
- NumPy: For handling numerical operations, especially in feature extraction.

Data Collection Process

The FracAtlas dataset, publicly available under a CC-BY 4.0 license, was collected from various medical institutions, ensuring a diverse range of fracture types and locations. Each X-ray image in the dataset underwent a thorough annotation process by medical professionals to accurately label the presence of fractures.

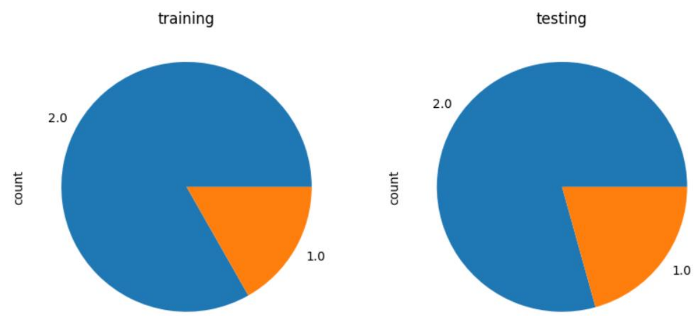

Figure 2: Spilitting Data

Canny Edge Segmentation

The Canny algorithm identifies the edges within an image by minimizing the error rate, detecting edges at different thresholds, and ensuring edge thinning. The process is defined by [15]:

$$E(x, y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial x}\right)^2} \tag{1}$$

Where $E(x, y)$ represents the edge intensity at point (x, y), and I is the image intensity. In **Figures 3** and **4** the results of image segmentation using canny features on the dataset are shown.

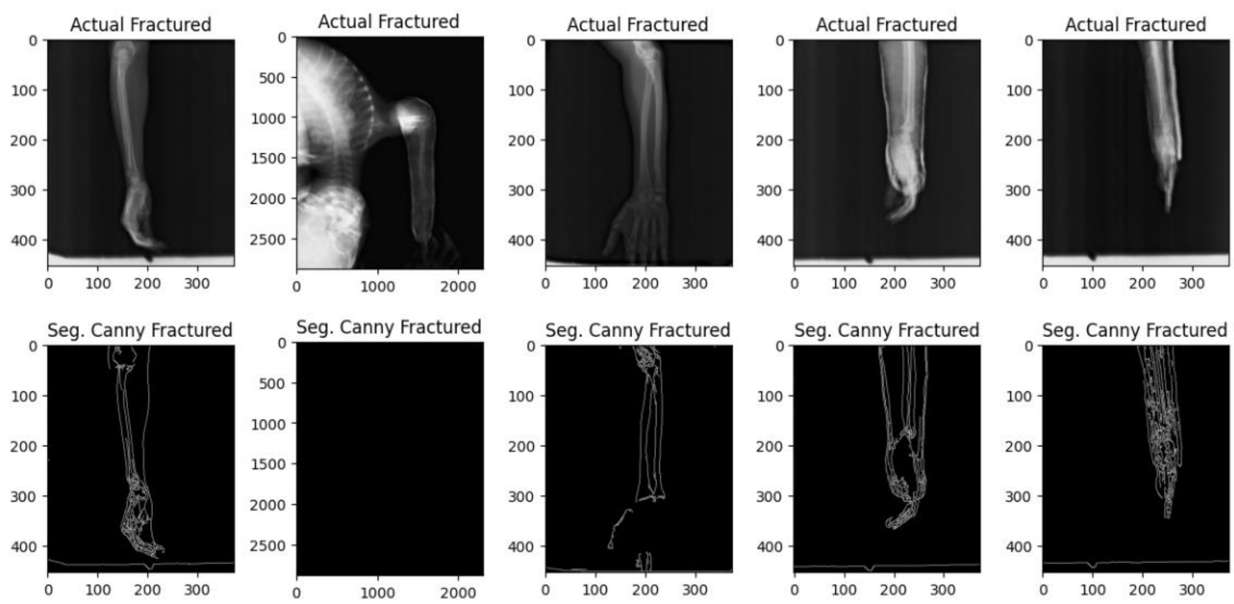

Figure 3: Canny Edges Detection Results for Fractured Class

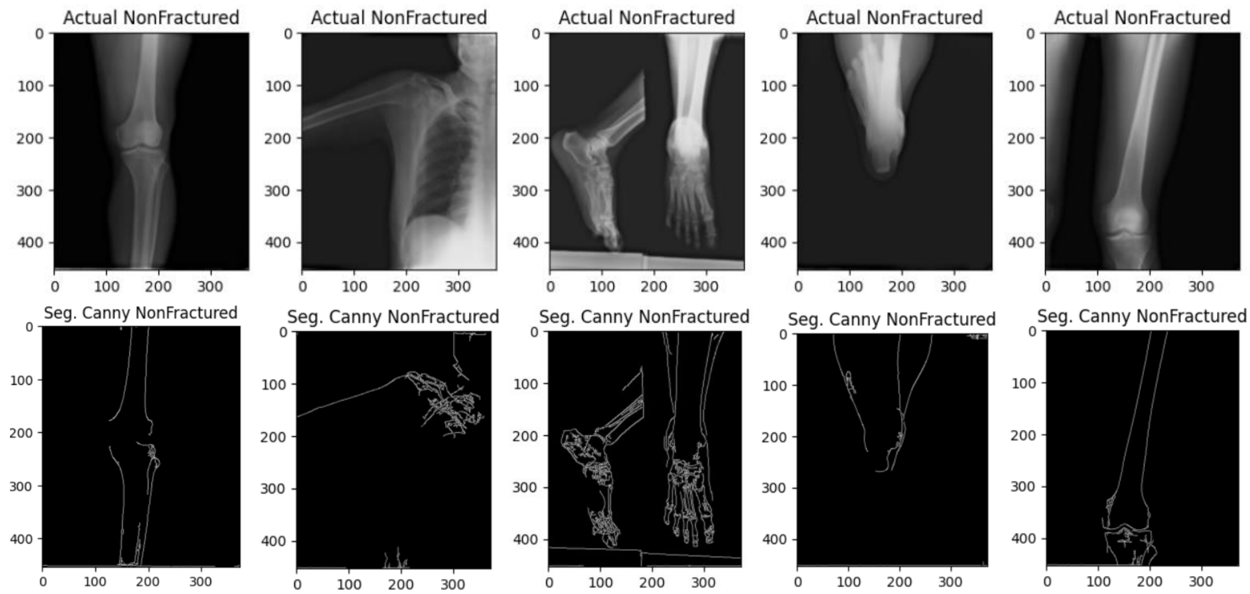

Figure 4: Canny Edges Detection Results for NonFractured Class

Feature Extraction using Hu Moments

Hu Moments are invariant to image transformations and provide a set of features representing the shape of objects within an image. The calculation of Hu Moments (H) from the moments (M) is as follows:

$$H_i = f(M) \tag{2}$$

Where i ranges from 1 to 7, representing the seven Hu Moments, and $f(M)$ denotes the specific calculation for each moment.

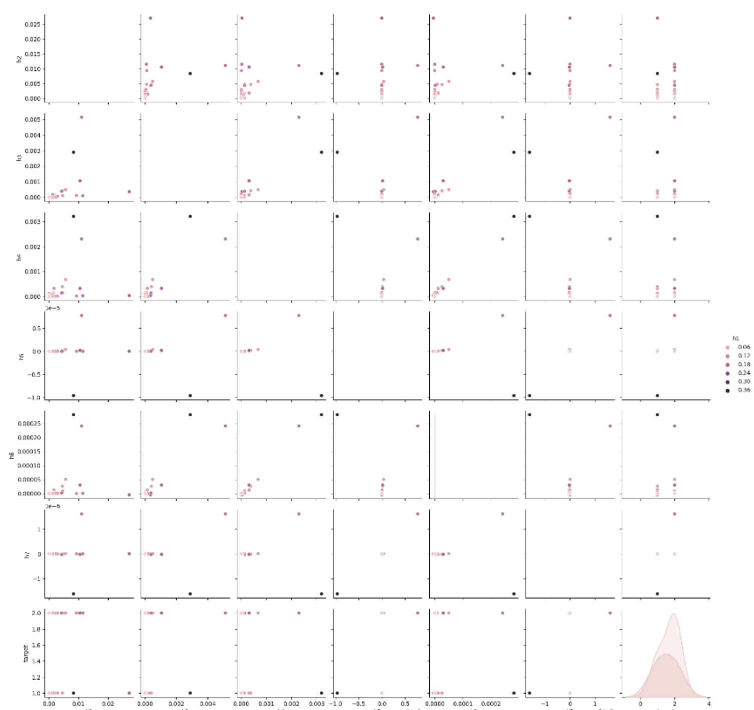

Figure 5: Scatter Plot Visualization of Extracted Hu Moments Features

Model Training and Testing

The Decision Tree model was trained using features extracted from the pre-processed images. The algorithm iteratively splits the data to maximize the purity of the target variable in each leaf node. The 5-fold cross-validation [16]–[19] process involves dividing the dataset into five parts, using four parts for training the model and the remaining part for testing, iterating this process five times. This method ensures that every data point is used for both training and testing, providing a comprehensive evaluation of the model's performance. The performance of the model was evaluated using accuracy, precision, recall, and F1-measure, calculated as follows [14], [20]–[22]:

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

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

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{4}$$

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

3. Result and Discussion

Our study focused on evaluating the efficacy of a Decision Tree classifier in diagnosing bone fractures from X-ray images, following pre-processing steps that included Canny edge detection for segmentation and Hu Moments for feature extraction. The classifier's performance was rigorously assessed using a 5-fold cross-validation approach, examining key metrics such as accuracy, precision, recall, and F1-score. The model's performance metrics across the 5 folds, providing a visual representation of the **Table 1**.

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

K-n	**Performa**			
	Accuracy	***Precision***	***Recall***	***F-Measure***
K-1	69.89%	72.70%	70.50%	71.56%
K-2	70.99%	72.27%	70.87%	71.52%
K-3	74.05%	72.77%	73.81%	73.27%
K-4	73.53%	73.58%	73.04%	73.31%
K-5	72.06%	73.25%	72.43%	72.84%
$\sum \boldsymbol{Avg}$	72.10%	72.91%	72.13%	72.50%

Table 1 provides a detailed view of the data processing and analysis revealed varying performance across the five folds, with accuracy rates ranging from approximately 69.89% to 74.05%. Precision values varied from 72.27% to 73.75%, recall rates were between 70.50% and 73.81%, and F1-scores spanned from 71.52% to 73.31%. These results indicate a moderate level of variability in the model's performance across different subsets of the dataset.

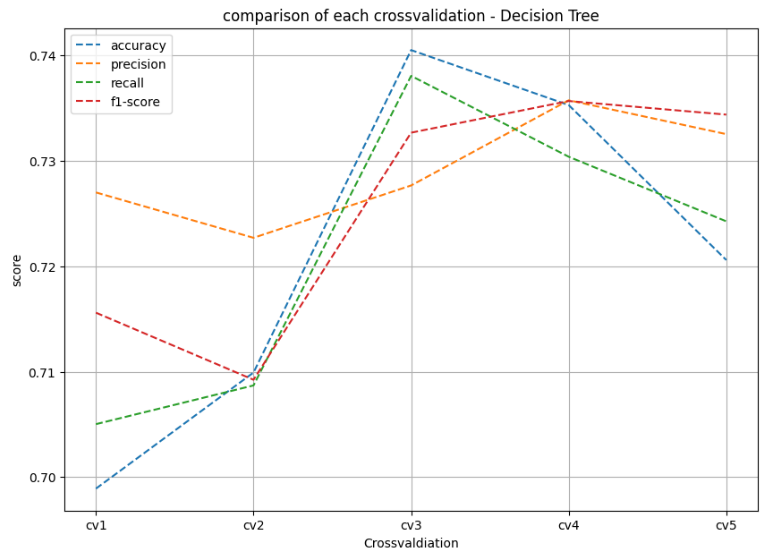

Figure 6: Performance Metrics Across 5-Fold Cross-Validation for the Decision Tree Algorithm

Figure 6 presented shown captures the comparative analysis of the Decision Tree model's performance across each fold of the 5-fold cross-validation process. This graph succinctly illustrates the scores of accuracy, precision, recall, and F1-score for each cross-validation fold, providing an immediate sense of the model's consistency and reliability in classifying bone fractures from X-ray images. As we examine the trends and patterns in the graph, we can draw insights into the strengths and potential areas of improvement for the predictive model.

The observed performance metrics suggest that the Decision Tree model, combined with the specified pre-processing techniques, can classify bone fractures in X-ray images with a reasonable degree of accuracy. However, the range of values across the folds highlights the model's variability in handling different data subsets. A notable finding is the balance between precision and recall, indicating the model's ability to maintain a relatively consistent trade-off between identifying true positives and minimizing false positives. This balance is crucial in medical diagnostics, where the cost of false negatives and positives can be significant.

Discussion

Interpretation and Evaluation of the Results

The Decision Tree model's performance, in light of the preprocessing steps applied, underscores the potential of machine learning in supporting medical diagnostics. The variability in performance metrics across folds underscores the importance of diverse and representative training data in developing robust diagnostic tools.

Relationship between the Research Results and Previous Research or Theory

Previous research has established the potential of machine learning algorithms in medical image analysis, often emphasizing the need for effective preprocessing to improve model performance. Our findings align with this literature, highlighting how specific techniques like Canny segmentation and Hu Moments can enhance the model's diagnostic capabilities.

Practical Implications of the Research Results

The application of a Decision Tree classifier for bone fracture diagnosis from X-ray images could significantly aid radiologists by providing a preliminary assessment tool, potentially improving diagnostic speed and accuracy. This could be particularly beneficial in high-volume clinical settings or areas with limited access to radiology expertise.

Limitations of the Research

One limitation of our study is the reliance on a singular dataset, which, despite its diversity, may not fully capture the complexity of real-world scenarios. Additionally, the Decision Tree model's interpretability comes at the cost of potentially lower performance compared to more complex models.

Recommendations for Further Research

Future research should explore the integration of more sophisticated machine learning models, such as deep learning, which could potentially improve upon the results obtained. Additionally, expanding the dataset and including more varied fracture types and imaging conditions could enhance the model's generalizability and accuracy. Further studies could also explore the impact of alternative pre-processing techniques and feature extraction methods on model performance.

This discussion offers insights into the potential of machine learning in enhancing diagnostic processes in the medical field, particularly in radiology. By building on these findings and addressing the identified limitations, future research can further the development of automated diagnostic tools, ultimately contributing to improved patient care.

4. Conclusion

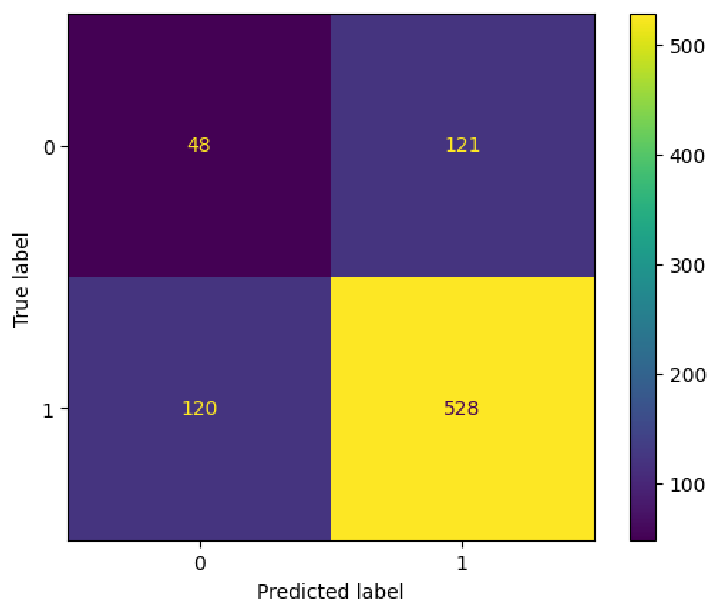

Figure 7: Confusion Matrix

Figure 7 depicted here provides a visual breakdown of the Decision Tree classifier's performance, highlighting the true positive, false positive, true negative, and false negative predictions made by the model. The colors ranging from deep purple to bright yellow represent the varying counts of predictions across each quadrant of the matrix, offering an intuitive understanding of the model's predictive accuracy for both fractured and non-fractured classes in the dataset.

In conclusion, the exploration of the Decision Tree classifier's application to the FracAtlas dataset has yielded informative insights into its utility for bone fracture detection in X-ray imaging. The classifier demonstrated moderate accuracy, precision, recall, and F1-scores through the 5-fold cross-validation process, with metrics indicative of its potential as a supportive diagnostic tool. The visualization of performance across folds and the resulting confusion matrix revealed the classifier's balanced capabilities in correctly identifying fractures versus non-fractures, albeit with some variability that suggests room for refinement. Our discussion has contextualized these findings within the broader scope of medical diagnostics, underlining the model's strengths in achieving a harmonious balance between sensitivity and specificity, which is crucial in clinical settings.

Addressing our initial research hypothesis, the data support the assertion that machine learning, particularly Decision Trees augmented by robust image processing techniques, can enhance the classification of bone fractures. This research contributes to the field of medical informatics by reinforcing the viability of such models in practical applications, with the potential to streamline workflows in radiological practices. Moving forward, it is recommended that subsequent investigations expand upon this foundation, possibly incorporating more advanced machine learning algorithms, larger and more varied datasets, and comprehensive pre-processing techniques. Such endeavours will not only refine diagnostic accuracy but also pave the way for integrating these models into real-world clinical environments, where they can aid in delivering expedient and reliable care to patients.

References:

[1] S. AbuRass, “Enhancing Convolutional Neural Network using Hu’s Moments,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 12, pp. 130–137, 2020, doi: 10.14569/IJACSA.2020.0111216.

[2] Y. Jusman, “Machine Learnings of Dental Caries Images based on Hu Moment Invariants Features,” *Proc. - 2021 Int. Semin. Appl. Technol. Inf. Commun. IT Oppor. Creat. Digit. Innov. Commun. within Glob. Pandemic, iSemantic 2021*, pp. 296–299, 2021, doi: 10.1109/iSemantic52711.2021.9573208.

[3] S. S. Gornale, “Automatic Detection and Classification of Knee Osteoarthritis Using Hu’s Invariant Moments,” *Front. Robot. AI*, vol. 7, 2020, doi: 10.3389/frobt.2020.591827.

[4] G. Sajiv, “Machine Learning based Analysis of Histopathological Images of Breast Cancer Classification using Decision Tree Classifier,” *6th Int. Conf. I-SMAC (IoT Soc. Mobile, Anal. Cloud), I-SMAC 2022 - Proc.*, pp. 989–995, 2022, doi: 10.1109/I-SMAC55078.2022.9987276.

[5] A. Anitha, “Disease prediction and knowledge extraction in banana crop cultivation using decision tree classifiers,” *Int. J. Bus. Intell. Data Min.*, vol. 20, no. 1, pp. 107–120, 2022, doi: 10.1504/IJBIDM.2022.119957.

[6] M. Aqib, “Classification of Edge Applications using Decision Tree, K-NN, & SVM Classifier,” *2022 IEEE Students Conf. Eng. Syst. SCES 2022*, 2022, doi: 10.1109/SCES55490.2022.9887690.

[7] K. A. N. Pelliza, “Analysis of the efficiency of the adaptive canny method for the detection of icebergs at open sea,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 3, pp. 459–464, 2020, doi: 10.5194/isprs-archives-XLII-3-W12-2020-459-2020.

[8] K. Benhamza, “Canny edge detector improvement using an intelligent ants routing,” *Evol. Syst.*, vol. 12, no. 2, pp. 397–406, 2021, doi: 10.1007/s12530-019-09299-0.

[9] Musa, “Automatic Face Mask Detection On Gates To Combat Spread Of Covid-19,” *Buana Inf. Technol. Comput. Sci. (BIT CS)*, vol. 3, no. 3, pp. 89–96, 2022, doi: 10.36805/bit-cs.v3i2.2759.

[10] 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.

[11] A. Jaya, “Analisis Sentimen pandangan public terhadap profesi PNS (Pegawai Negeri Sipil) dari Twiter menerapkan Indonesian,” *Indones. J. Data Sci.*, vol. 4, no. 1, pp. 38–44, 2023.

[12] R. Rohan, “Classification of cardiac arrhythmia diseases from obstructive sleep apnea signals using decision tree classifier,” *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, vol. 12, pp. 248–264, 2020.

[13] I. A. P. Banlawe, “Decision Tree Learning Algorithm and Naïve Bayes Classifier Algorithm Comparative Classification for Mango Pulp Weevil Mating Activity,” *2021 IEEE Int. Conf. Autom. Control Intell. Syst. I2CACIS 2021 - Proc.*, pp. 317–322, 2021, doi: 10.1109/I2CACIS52118.2021.9495863.

[14] 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.

[15] Z. Hu, “Canny Algorithm Enabling Precise Offline Line Edge Roughness Acquisition in High-Resolution Lithography,” *ACS Omega*, 2022, doi: 10.1021/acsomega.2c06769.

[16] Z. Xiong, “Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation,” *Comput. Mater. Sci.*, vol. 171, 2020, doi: 10.1016/j.commatsci.2019.109203.

[17] O. Karal, “Performance comparison of different kernel functions in SVM for different k value in k-fold cross-validation,” *Proc. - 2020 Innov. Intell. Syst. Appl. Conf. ASYU 2020*, 2020, doi: 10.1109/ASYU50717.2020.9259880.

[18] S. Ortiz-Toquero, “Classification of Keratoconus Based on Anterior Corneal High-order Aberrations: A Cross-validation Study,” *Optom. Vis. Sci.*, vol. 97, no. 3, pp. 169–177, 2020, doi: 10.1097/OPX.0000000000001489.

[19] K. M. Bain, “Cross-validation of three Advanced Clinical Solutions performance validity tests: Examining combinations of measures to maximize classification of invalid performance,” *Appl. Neuropsychol.*, vol. 28, no. 1, pp. 24–34, 2021, doi: 10.1080/23279095.2019.1585352.

[20] 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, [Online]. Available: https://jurnal.yoctobrain.org/index.php/ijodas/article/view/54.

[21] D. Anggreani, I. A. E. Zaeni, A. N. Handayani, H. Azis, and A. R. Manga’, “Multivariate Data Model Prediction Analysis Using Backpropagation Neural Network Method,” in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, 2021, pp. 239–243, doi:

10.1109/EIConCIT50028.2021.9431879.

[22] 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.