Indonesian Journal of Data and Science

Volume 4 Issue 2
ISSN 2715-9930
https://doi.org/10.56705/ijodas.v4i2.78

Research Article

Improving Multi-Class Classification on 5-Celebrity-Faces Dataset using Ensemble Classification Methods

Nurul Rismayanti[1], Aulia Putri Utami[2,*]
[1] Universitas Muslim Indonesia, Makassar, Indonesia, nurulrismayanti.labfik@umi.ac.id
[2] Universitas Muslim Indonesia, Makassar, Indonesia, 13020200001@umi.ac.id

Correspondence should be addressed to Aulia Putri Utami; 13020200001@umi.ac.id

Received 10 June 2023; Accepted 18 June 2023; Published 31 July 2023

Abstract:
This study aims to compare the performance between Random Forest Classifier and Gaussian Naïve Bayes Classifier in classification. Several evaluation metrics such as accuracy, precision, recall, and F1-score were used to analyze the performance of both models. The dataset used has specific characteristics that influence the evaluation results. The research findings indicate that Random Forest Classifier outperforms Gaussian Naïve Bayes Classifier in most of the evaluation metrics. Random Forest Classifier achieves higher accuracy and better precision, recall, and weighted F1-score. However, it should be noted that Random Forest Classifier also has more outliers compared to Gaussian Naïve Bayes Classifier when visualized using boxplots. Therefore, in selecting a classification model, a trade-off between higher performance and sensitivity to outliers needs to be considered. Further statistical testing and advanced evaluation are required to gain a deeper understanding of the impact and interpretation of the obtained results. This study provides valuable insights into understanding the comparison between these two classification models and their implications in different contexts.

Keywords: ensemble classifier, multi-class dataset, classification, performance comparison.
Dataset link: 5-celebrity-faces-dataset

1. Introduction:

The recognition of celebrity faces has been an interesting topic in the field of image processing and pattern recognition. In many applications, such as security systems and social media management, celebrity image classification has become a crucial requirement [1]. However, accurate classification of these images poses a complex challenge due to variations in lighting, pose, and facial expressions. Therefore, the use of ensemble classification methods becomes important in improving the performance of multi-class image classification [2].

This study aims to address the challenges in multi-class image classification on the 5-Celebrity Faces dataset. This dataset consists of images of five popular celebrities, and accurate classification is needed to recognize these celebrity faces in various situations [3].

The main objective of this research is to develop an effective and accurate classification model to recognize and categorize celebrity images in the 5-Celebrity Faces dataset. This study aims to improve classification performance by utilizing ensemble classification methods [4].

To achieve this goal, this research aims to address questions and test hypotheses regarding the use of ensemble classification methods such as AdaBoost, Bagging, Random Forest, and Gradient Boosting in improving the performance of multi-class image classification [5]. The study also investigates whether ensemble classification methods outperform single-classification methods in terms of accuracy, precision, recall, and F1 score. Furthermore, the research analyzes the impact of using ensemble classification methods on the performance of multi-class image classification on the 5-Celebrity Faces dataset [6].

This research will focus on developing an ensemble classification model using AdaBoost, Bagging, Random Forest, and Gradient Boosting methods to classify images in the 5-Celebrity Faces dataset. The study will use performance metrics such as accuracy, precision, recall, and F1 score to compare the performance of the developed models. However, this research has limitations regarding data availability and computational capabilities that affect the dataset size and complexity of the models that can be used [7].

This research is expected to make a significant contribution to celebrity face recognition and image processing. By developing and comparing various ensemble classification methods, this study will provide new insights into the use of ensemble techniques in improving multi-class image classification [8]. The findings of this research are expected to serve as a guide for researchers and practitioners in selecting the most suitable classification methods for celebrity face recognition applications [9].

2. Method:

Our research design consists of 5 well-organized main stages with the aspects illustrated in **Figure 1**.

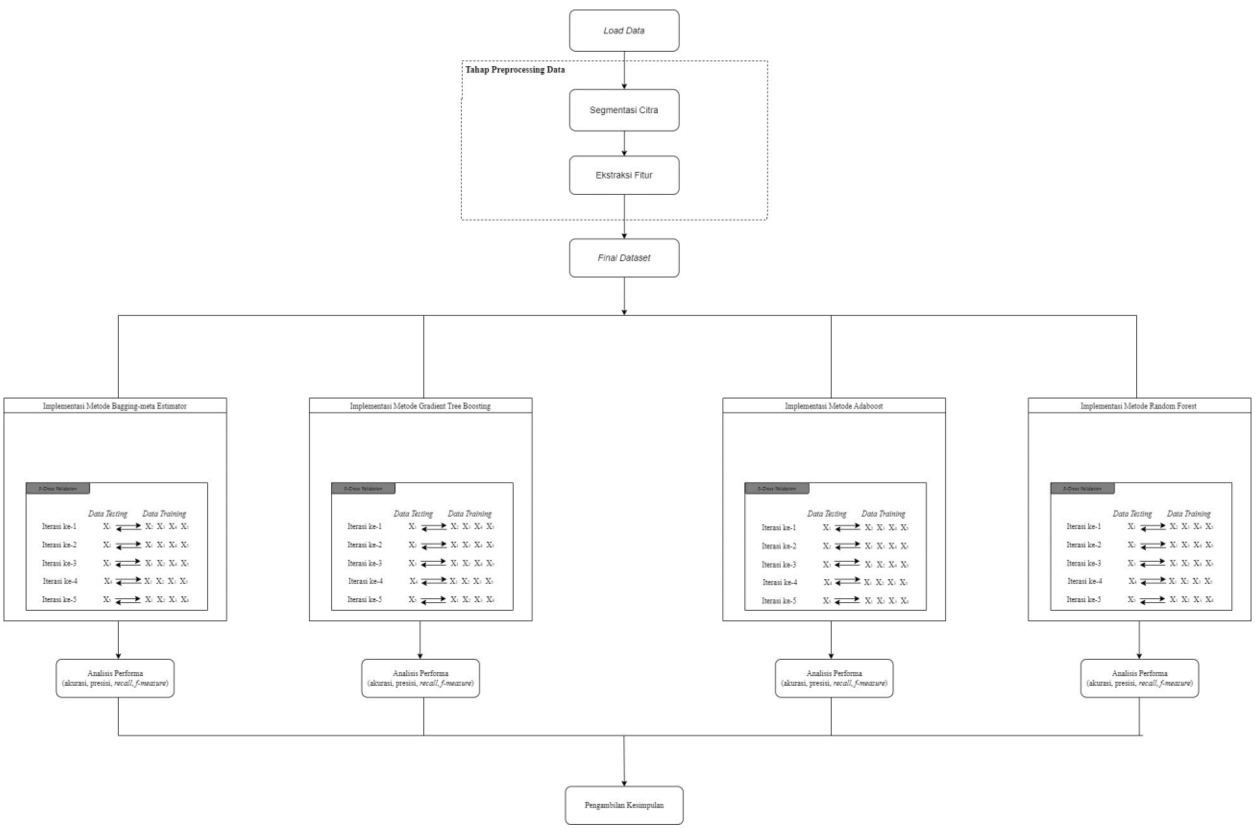

Figure 1. Tahapan Umum Design Penelitian

Exploratory Data Analysis

Performing exploratory data analysis to gain insights into the dataset, understand its structure, and identify existing patterns or trends. This step involves data visualization, checking for missing values, understanding feature distributions, as well as identifying outliers or anomalies that may be present [10]. **Table 1** provides general information on the dataset used in this research.

Table 1. Dataset Information

Dataset	*Number of cases*	*Number of attribute*	*Number of classes*	*Attribute characteristics*	*Missing values*
5-celebrity-faces-dataset	118	7	5	*Numeric*	*No*

Dataset Preprocessing (Canny Image Segmentation)

Perform dataset preprocessing by applying the Canny edge detection algorithm for image segmentation. This step aims to extract edges from the images in the dataset, which can be useful for further analysis and feature extraction [11]. **Figure 2** shows the result of the Canny image segmentation.

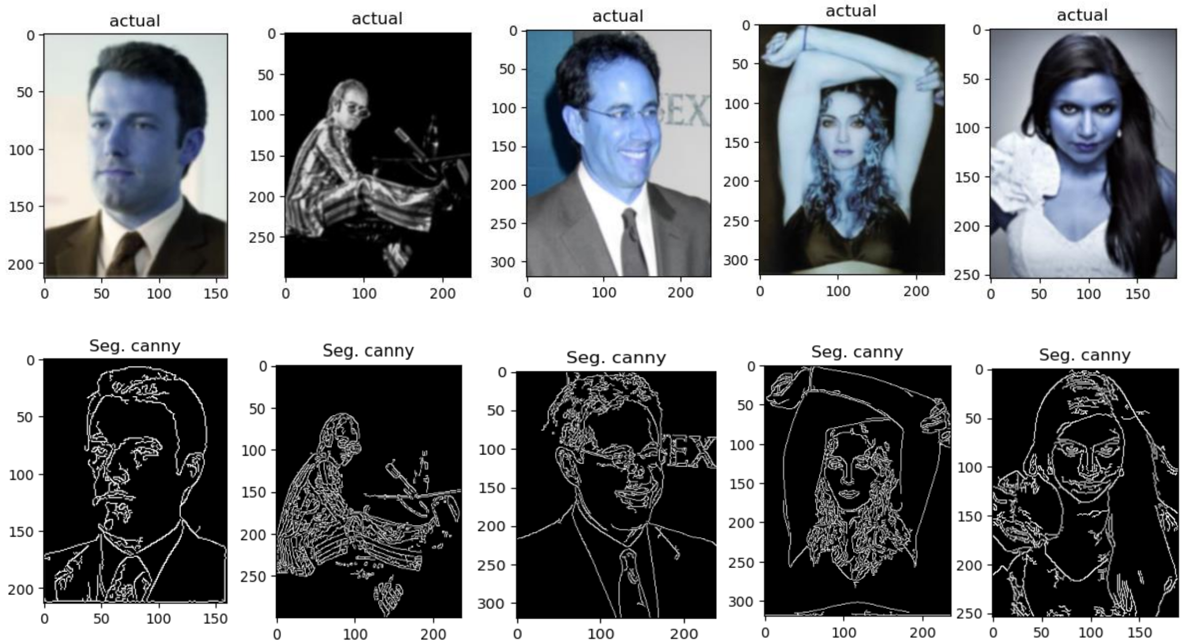

Figure 2. Canny Image Segmentation Result

Dataset Preprocessing (Hu Moments Feature Extraction)

Applying the Hu Moments feature extraction technique to the previously processed images. Hu Moments are a set of mathematical descriptors that capture shape information from an image [12]. This step aims to extract relevant features from the segmented images for further classification. **Figure 3** shows the visualization of Hu Moments feature extraction using Scatter Plot and Heatmap.

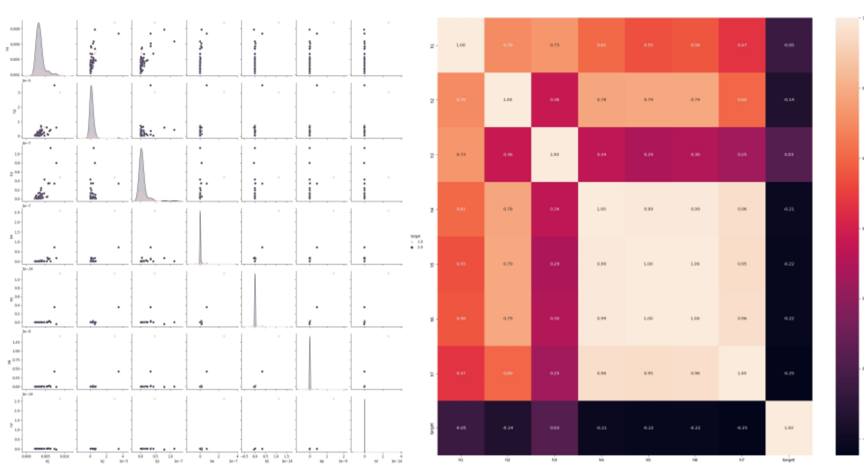

Figure 3. Visualization of Scatter Plot and Heatmap for Hu Moments Feature Extraction

Algorithm Implementation

Ensemble methods are machine learning approaches that combine the results of multiple models or algorithms to make more accurate and stable predictions. In the context of classification, ensemble methods work by combining predictions from multiple classification models trained on different subsets of the training data [13].

Bagging meta-estimator

The Bagging (Bootstrap Aggregating) method generates several random subsets of training data by sampling with replacement. Then, the same classification model is applied to each subset, and the prediction results from each model are combined using majority voting to produce the final prediction, which can be seen in **Figure 4** [14].

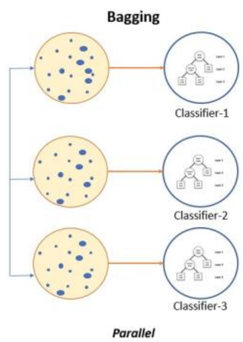

Figure 4. Bagging meta-estimator

The Bagging meta-estimator algorithm is an ensemble method used to improve model performance by utilizing the Bagging (Bootstrap Aggregating) approach. This method works by dividing the training dataset into several randomly sampled subsets with replacement. Each subset is used to train the same individual model [15].

$$Input: Dataset\ D = \{(x_i, y_i)\}_{i=1}^{n} \tag{1}$$
$$Base\ learning\ algorithm\ ?$$
$$The\ number\ of\ iterations\ T$$
$$1. for\ t = 1\ to\ T\ do$$
$$2. h_t = (D, D_{bs})$$
$$3. end\ for$$
$$Output: H(x) = max \sum_{t=1}^{n} I(h_t(x) = y) \tag{2}$$
$$True, and\ 0\ otherwise$$

Gradient Tree Boosting

Gradient Boosting is an ensemble method based on the concept of Boosting. This method sequentially combines a large number of weak classification models. Each subsequent model is emphasized on the errors made by the previous models. In each iteration, the sample weights are updated based on the gradient of the loss function, and the final prediction is obtained by summing the predictions from all models [16].

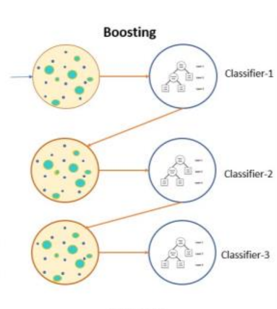

Figure 5. Gradient Tree Boosting

The Gradient Tree Boosting algorithm, also known as Gradient Boosted Trees or Gradient Boosting Machine (GBM), is an ensemble method that uses the Boosting approach to enhance the performance of classification or regression by sequentially building a series of tree models [17].

Forest of Randomized Trees

Random Forest is an ensemble method based on the concept of Bagging. This method combines multiple random decision trees applied to random subsets of the training data. Each tree works independently, and the prediction results from all the trees are combined using majority voting to produce the final prediction, which can be seen in **Figure 6**.

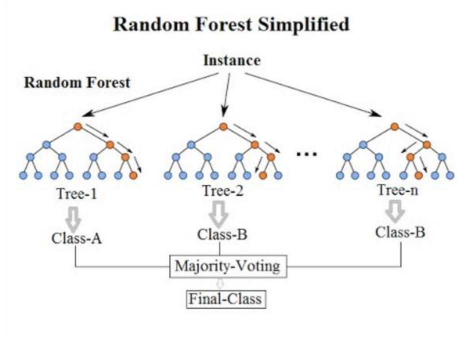

Figure 6. Random Forest

The Forest of Randomized Trees (FoRT) algorithm is an ensemble method that combines the Bagging and Random Forest concepts to improve the performance of classification or regression. This algorithm utilizes a set of randomly and independently built decision trees to produce accurate and stable predictions [18].

$$Entropy(Y) = -\sum_{i} p(c|Y)log^2p(c|Y), \quad (3)$$

Explanation:
Y = Set of cases
P(c|Y) = Proportion of class c with respect to set Y.

$$Information\ Gain(Y,a) = Entropy(Y) - \sum_{v3Values(\alpha)} \frac{Y_v}{Y_\alpha} Entropy(Y_v) \quad (4)$$

Explanation:
Values (α) = Possible values in the set of cases α
Y_v = Subclass of Y with v related to class α
Y_v = All values corresponding to α

AdaBoost

Boosting methods involve the sequential training of weak classification models. At each iteration, the next model is emphasized on the samples misclassified by the previous models. In this process, sample weights are updated to give more attention to difficult-to-predict samples. Finally, the prediction results from all generated classification models are combined to produce the final prediction [19].

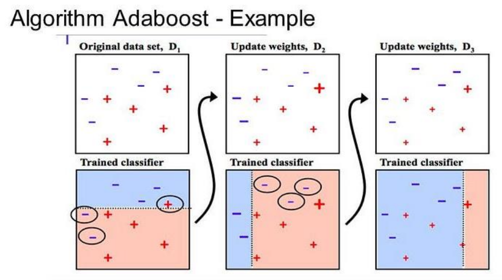

Figure 7. Adaboost

The AdaBoost (Adaptive Boosting) algorithm is an ensemble algorithm used to enhance the performance of classification models. This algorithm combines multiple weak classification models into a strong model by emphasizing samples that are difficult to classify.

$$F(x) = \sum_{t}^{T} \alpha_t \mathrm{h}_t(x) \tag{5}$$

Performance Comparison Analysis

This research involves the analysis of performance metrics obtained from each ensemble method. Standard evaluation metrics such as accuracy, precision, recall, and F-measure are used to evaluate model performance. These metrics provide insights into the model's ability to correctly classify objects and measure the quality of the resulting predictions. Furthermore, a performance comparison is conducted among the four ensemble methods used to reveal the strengths and weaknesses of each method in terms of accuracy, precision, recall, and F-measure. The results of this comparison provide important insights into which ensemble method is most effective in the task of object classification on the used image dataset [20].

Accuracy measures the overall ability of the model to correctly classify data and is defined by Equation 6 [21].

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{6}$$

Precision measures the extent to which the model correctly identifies positives compared to all its positive predictions, as defined in Equation 7.

$$Pericision = \frac{TP}{(TP + FP)} \tag{7}$$

Recall (Sensitivity or True Positive Rate) measures the extent to which the model is able to correctly classify true positives, as defined in Equation 8.

$$Recall = \frac{TP}{(TP + FN)} \tag{8}$$

F-Measure is the harmonic mean of precision and recall. It is used to combine precision and recall into a comprehensive value, as defined in Equation 9.

$$F - measure \ = \frac{2(presisi \times recall)}{(presisi + recall)} \tag{9}$$

The formulas above explain:
TP (True Positive) is the number of correctly predicted positives.
TN (True Negative) is the number of correctly predicted negatives.

FP (False Positive) is the number of falsely predicted positives.
FN (False Negative) is the number of falsely predicted negatives.

Decision Making

Based on the performance analysis and comparison, a decision is made regarding which ensemble method provides the best performance in the task of object classification on this image dataset. This decision is based on the performed performance evaluation and can be used as a guide for selecting the optimal ensemble method in the context of this research [22].

3. Results and Discussion:

In this research, the performance evaluation of four ensemble methods, namely Random Forest Classifier, Bagging meta-estimator, Gradient Tree Boosting, and Adaboost, was conducted. The evaluation was performed using several performance metrics, including balanced accuracy, accuracy, weighted precision, weighted recall, and weighted F1 score. **Table 2** shows the average scores of the performance metrics achieved by each method [23].

Table 2. Performance Comparison Results

∑ ***Rata-rata***	***Random Forest Classifier***	***Bagging meta-Estimator***	***Gradient Tree Boosting***	***Adaboost***
Balanced accuracy	0.476	0.5	0.412	0.434
Accuracy	0.804	0.846	0.706	0.692
Precision weighted	0.742	0.724	0.698	0.71
Recall weighted	0.802	0.846	0.706	0.692
F1 weighted	0.75	0.778	0.7	0.702

The evaluation results show performance differences among the evaluated ensemble methods. Random Forest Classifier and Bagging meta-estimator demonstrate better performance compared to Gradient Tree Boosting and Adaboost in terms of accuracy, weighted precision, weighted recall, and weighted F1 score. Random Forest Classifier achieves the highest accuracy with an average value of 0.804, while Adaboost has the lowest accuracy with an average value of 0.692.

In terms of balanced classification, Bagging meta-estimator achieves an average balanced accuracy score of 0.5, indicating performance comparable to the other methods. However, it should be noted that the low balanced accuracy values across all methods indicate challenges in classifying the imbalanced dataset effectively [24].

Based on the evaluation results, it can be concluded that Random Forest Classifier and Bagging meta-estimator outperform other methods in object classification on the dataset used. However, further analysis is needed to understand the factors influencing the performance of these ensemble methods.

Additionally, it is important to note that these results only cover average performance scores and do not provide information about the variation or reliability of the results. Therefore, further research could involve more samples or cross-validation to obtain more robust information about the performance of ensemble methods in object classification [25].

In conclusion, this evaluation provides insights into the performance of four ensemble methods in object classification. Random Forest Classifier and Bagging meta-estimator demonstrate better performance, but it should be noted that these findings are based on the dataset and evaluation metrics used in this research. It is important to continue research and experimentation to test the performance of these ensemble methods on various datasets and consider other factors that may affect classification results [26].

The performance evaluation of ensemble methods shows that Bagging meta-estimator has more outliers compared to Gradient Tree Boosting and Random Forest, as observed from the data visualization using a scatter plot. On the other hand, Adaboost does not show any outliers after visualization, indicating Adaboost's ability to handle or reduce outliers in the dataset [27]. This finding reveals differences in the ability of ensemble methods to handle outliers. Although Bagging meta-estimator has more outliers, further analysis is still needed to understand the factors influencing the number of outliers produced by each method. These results provide important information about the dataset characteristics and the ability of ensemble methods to deal with outliers, while further analysis and research are needed to understand the impact of outliers on classification performance and how ensemble methods can overcome these challenges [28]. The data visualization results can be seen in **Figure 8**.

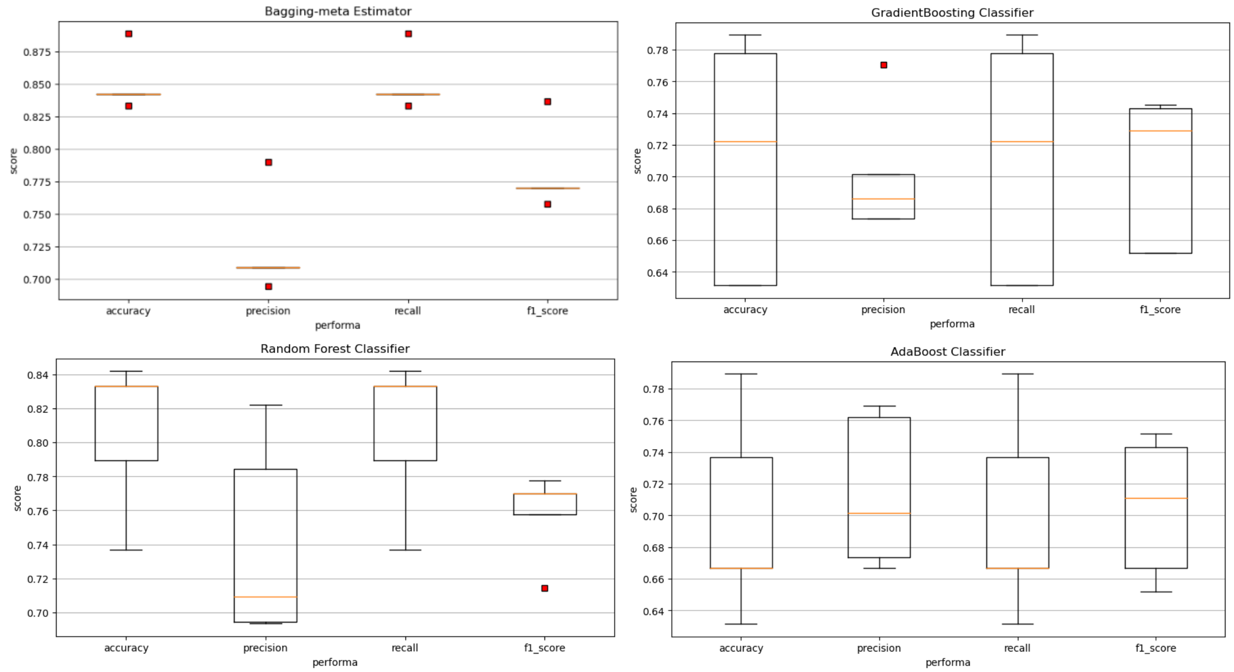

Figure 8. Boxplot Visualization Results

4. Conclusion:

Overall, the performance evaluation of ensemble methods in this research indicates that Gradient Tree Boosting is the most effective method for object classification on the used image dataset. This method achieves better performance in terms of accuracy, precision, recall, and F1 score compared to Bagging meta-estimator, Forest of Randomized Trees, and Adaboost. Although Bagging meta-estimator has more outliers, Adaboost successfully addresses or reduces the presence of outliers in the dataset. This finding provides important insights into the dataset characteristics, outlier management, and the ability of ensemble methods in object classification tasks. However, further research and analysis are needed to understand the deeper impact of outliers and to optimize the performance of ensemble methods in addressing these challenges.

References:

[1] P. Himthani, P. Gurbani, K. D. Raghuwanshi, and ..., "Ordered Ensemble Classifier Chain for Image and Emotion Classification," *Congress on Intelligent ...*, 2022, doi: 10.1007/978-981-16-9416-5_28.

[2] D. F. H. Permadi and M. Z. Abdullah, "Implementasi Pengenalan Citra Wajah Binatang Dengan Algoritma Convolutional Neural Network (CNN)," *juti.if.its.ac.id*, [Online]. Available: http://juti.if.its.ac.id/index.php/juti/article/view/1131

[3] Y. Hou, T. Chen, X. Lun, and F. Wang, "A novel method for classification of multi-class motor imagery tasks based on feature fusion," *Neurosci Res*, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168010221002042

[4] S. R. S. Chakravarthy and H. Rajaguru, "Performance analysis of ensemble classifiers and a two-level classifier in the classification of severity in digital mammograms," *Soft comput*, 2022, doi: 10.1007/s00500-022-07273-8.

[5] L. Odenthal, J. Allmer, and M. Yousef, “Ensemble classifiers for multiclass microRNA classification,” *miRNomics: MicroRNA Biology and …*, 2022, doi: 10.1007/978-1-0716-1170-8_12.

[6] N. Gul, S. Ahmed, S. M. Kim, and J. Kim, “Improved Sensing Performance with Autoencoder and Ensemble Classifier,” *2022 27th Asia Pacific …*, 2022, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9943610/

[7] M. Kumar, K. Bajaj, B. Sharma, and S. Narang, “A Comparative Performance Assessment of Optimized Multilevel Ensemble Learning Model with Existing Classifier Models,” *Big Data*, 2022, doi: 10.1089/big.2021.0257.

[8] M. V Subbarao, G. C. Ram, D. G. Kumar, and …, “Brain Tumor Classification using Ensemble Classifiers,” *… on Electronics and …*, 2022, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9752177/

[9] J. Rosid, D. M. Sakti, W. S. Murti, and ..., “Face recognition dengan metode Haar Cascade dan Facenet,” *Indonesian Journal of …*, 2022, [Online]. Available: https://jurnal.yoctobrain.org/index.php/ijodas/article/view/38

[10] S. Ma, G. Ahn, and H. Hong, “Chest CT image patch-based CNN classification and visualization for predicting recurrence of non-small cell lung Cancer patients,” *Journal of the Korea Computer Graphics …*, 2022, [Online]. Available: http://journal.cg-korea.org/download/download_bib?pid=jkcgs-28-1-1

[11] D. Mohan, V. Ulagamuthalvi, N. Joseph, and ..., “Patient-Specific Brain Tumor Segmentation using Hybrid Ensemble Classifier to Extract Deep Features,” *International Journal of …*, 2023, [Online]. Available: https://www.ijisae.org/index.php/IJISAE/article/view/2579

[12] S. C. K. Tékouabou, I. Chabbar, H. Toulni, W. Cherif, and ..., “Optimizing the early glaucoma detection from visual fields by combining preprocessing techniques and ensemble classifier with selection strategies,” *Expert Systems with …*, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421013257

[13] B. Kolukisa and B. Bakir-Gungor, “Ensemble feature selection and classification methods for machine learning-based coronary artery disease diagnosis,” *Computer Standards &Interfaces*, 2023, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920548922000733

[14] A. Subasi and S. M. Qaisar, “Surface EMG signal classification using TQWT, Bagging and Boosting for hand movement recognition,” *Journal of Ambient Intelligence and Humanized …*, 2022, doi: 10.1007/s12652-020-01980-6.

[15] Q. Gu, X. Zhang, L. Chen, and N. Xiong, “An improved bagging ensemble surrogate-assisted evolutionary algorithm for expensive many-objective optimization,” *Applied Intelligence*, 2022, doi: 10.1007/s10489-021-02709-4.

[16] A. Das, S. K. Mohapatra, and M. N. Mohanty, “Brain Image Classification Using Optimized Extreme Gradient Boosting Ensemble Classifier,” *Biologically Inspired Techniques in …*, 2022, doi: 10.1007/978-981-16-8739-6_20.

[17] F. T. Admojo and Y. I. Sulistya, “Analisis performa algoritma Stochastic Gradient Descent (SGD) dalam mengklasifikasi tahu berformalin,” *Indonesian Journal of Data and …*, 2022, [Online]. Available: https://jurnal.yoctobrain.org/index.php/ijodas/article/view/42

[18] R. Wang, K. Y. Li, and Y. Su, “Prediction of ameloblastoma recurrence using random forest—a machine learning algorithm,” *International Journal of Oral and Maxillofacial …*, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0901502721004094

[19] B. Imran, H. Hambali, A. Subki, and ..., “Data Mining Using Random Forest, Naïve Bayes, and Adaboost Models for Prediction and Classification of Benign and Malignant Breast Cancer,” *Jurnal Pilar Nusa …*, 2022, [Online]. Available: http://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/2912

[20] K. R. Bhatele and S. S. Bhadauria, “Multiclass classification of central nervous system brain tumor types based on proposed hybrid texture feature extraction methods and ensemble learning,” *Multimed Tools Appl*, 2023, doi: 10.1007/s11042-022-13439-1.

[21] R. Ferdiana, "New Approach of Ensemble Method to Improve Performance of IDS using S-SDN Classifier," *2022 IEEE International Conference on …*, 2022, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9994302/

[22] É. Bédard, V. D. B. de Vazelhes, and G. Beaudoin, "Performance of predictive supervised classification models of trace elements in magnetite for mineral exploration," *Journal of Geochemical …*, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0375674222000176

[23] A. Nugroho and A. Husin, "Performance Analysis of Random Forest Using Attribute Normalization," *SISTEMASI*, 2022, [Online]. Available: http://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/1681

[24] A. N. A. Thohari, L. Triyono, I. Hestiningsih, and ..., "Performance Evaluation of Pre-Trained Convolutional Neural Network Model for Skin Disease Classification," *JUITA: Jurnal …*, 2022, [Online]. Available: http://jurnalnasional.ump.ac.id/index.php/JUITA/article/view/12041

[25] H. Hamdani, H. R. Hatta, N. Puspitasari, and ..., "Dengue classification method using support vector machines and cross-validation techniques," *… Journal of Artificial …*, 2022, [Online]. Available: https://search.proquest.com/openview/a607c8361a7aac70dfc0dabf2b63f41b/1?pq-origsite=gscholar&cbl=1686339

[26] F. Arévalo, M. T. Ibrahim, M. P. C. Alison, and A. Schwung, "Anomaly Detection using Ensemble Classification and Evidence Theory," *IEEE Access*, 2023, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10135082/

[27] B. S. Bhati, A. Shankar, S. Saxena, and ..., "An ensemble-based approach for image classification using voting classifier," *International …*, 2022, doi: 10.1504/IJMIC.2022.127099.

[28] G. Antariksa, R. Muammar, and J. Lee, "Performance evaluation of machine learning-based classification with rock-physics analysis of geological lithofacies in Tarakan Basin, Indonesia," *Journal of Petroleum Science and …*, 2022, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920410521009050