Jurnal Teknik Informatika C.I.T Medicom, 15 (1) (2023) 48-57

Published by: Institute of Computer Science (IOCS)

Jurnal Teknik Informatika C.I.T Medicom

Journal homepage: www.medikom.iocspublisher.org

Compares the effectiveness of the bagging method in classifying spices using the histogram of oriented gradient feature extraction technique

Muhathir

Teknik Informatika, Fakultas Teknik, Universitas Medan Area, Indonesia

Article Info

Article history:

Received: Mar 16, 2023
Revised: Mar 20, 2023
Accepted: Mar 30, 2023

Kata Kunci:

Bagging Method;
Histogram of Oriented Gradient;
Spices.

Abstract

Spice classification is a crucial task in the food industry to ensure food safety and quality. This study focuses on the classification of spices using the Histogram of Oriented Gradient (HoG) feature extraction method and bagging method. The objective of this research is to compare the performance of three different models of bagging method, including Bootstrap Aggregating (Bagging), Random Forests, and Extra Tree Classifier, in classifying spices. The evaluation metrics used in this research are Precision, Recall, F1-Score, F2-Score, Jaccard Score, and Accuracy. The results show that the Random Forest model achieved the best performance, with precision, recall, F1-score, F2-Score, Jaccard, and Accuracy values of 0.861, 0.8633, 0.8587, 0.8607, 0.7694, and 0.8733 respectively. On the other hand, the Extra Tree Classifier had the lowest performance with precision, recall, F1-score, F2-Score, Jaccard, and Accuracy values of 0.7034, 0.7958, 0.7037, 0.7047, 0.5635, and 0.72 respectively. Overall, the results indicate a fairly good success rate in classifying spices using the HoG feature extraction method and bagging method. However, further evaluation is needed to improve the accuracy of the classification results, such as increasing the number of training data or considering the use of other feature extraction methods. The findings of this research may have significant implications for the food industry in ensuring the quality and safety of food products.

Corresponding Author:

Muhathir,
Teknik Informatika, Fakultas Teknik,
Universitas Medan Area, Indonesia
Jl. Setia Budi No.79 B, Kota Medan, Sumatera Utara,Indonesia 20112
Email: muhathir@staff.uma.ac.id

1. Introduction

Indonesia has gained a reputation for producing a wide variety of spices, which has attracted the attention of European countries in the past. Despite this history, Indonesia continues to be a significant player in the global spice market, accounting for 21.06% of global spice production in 2013. This presents a great opportunity for businesses in the plantation subsector to sell Indonesian spices globally, which could greatly benefit the country's economy[1], [2]. Furthermore, the demand for natural ingredients like spices is expected to continue to grow due to factors such as population growth, health concerns, and environmental sustainability. Spices have been used for thousands of years as flavor enhancers, colorings, and aromas and are also rich in antioxidants, which can help protect against disease[3].

Spices have been utilized for centuries due to their appealing fragrance and taste-enhancing properties [4]. They are considered a unique type of food additive that has been used for flavoring, coloring, and preserving food. Additionally, spices are recognized for their medicinal benefits and have been utilized in traditional medicine practices for a long time [5], [6]. Some examples of spices that come from seed plants are cumin, nutmeg, pepper, fennel, and coriander. Meanwhile, spices that come from rhizomes include turmeric, ginger, galangal, and fingerroot [7].

According to a study conducted at SMKN 9 Bandung [8], almost half of the students (47%) were not familiar with spices and herbs. With advancements in digital image processing technology, it is now possible to automatically classify spices and herbs using image classification techniques. The objective of image classification is to imitate human ability in comprehending digital image information so that computers can classify objects depicted in images in the same manner as humans. There are various techniques available to solve this problem, and one such method is Naïve Bayes Classification Method and HOG (Histogram of Oriented Gradients) Feature Extraction, as discussed in a study[9] on fruit classification.

The Histogram of Gradient (HOG) algorithm is a method used for detecting objects by extracting distinctive features based on gradient information from each pixel. This method involves extracting HOG features from multiple window sizes of an image, and it has been widely studied in the field of computer vision. The HOG feature extraction algorithm has proven to be effective in various computer vision applications, as highlighted in several studies [10]–[13]. In a study by [14]–[16] HOG features were used for various computer vision tasks and achieved impressive results. Recently, Leidiyana & Warta [17] used the HOG feature extraction method along with the SVM algorithm for fruit classification.

Bagging (Bootstrap Aggregating) is an ensemble learning algorithm used to improve the performance of machine learning models by combining several simple models. Bagging or Bootstrap Aggregating is an ensemble learning algorithm first introduced by Leo Breiman in 1996[18]. This algorithm works by creating several different prediction models using random samples of training data taken with replacement (bootstrap). Next, a more robust model is created by utilizing a voting or averaging approach to make decisions based on the predictions of all individual models that were previously combined. Random Forest is a machine learning algorithm used for classification, regression, and other tasks. This algorithm works by combining many decision trees created randomly. Each tree in the Random Forest is created based on a random subsample of training data and a random subset of available features. Then, the final prediction is made by combining the prediction results of all trees in the ensemble[19]. ExtraTreeClassifier, or Extremely Randomized Trees, is an ensemble learning method for classification and regression tasks, which uses a randomized decision tree algorithm to create a forest of decision trees. It is similar to the Random Forest algorithm, but with a key difference: ExtraTreeClassifier selects a random subset of features for each split in the decision tree, rather than evaluating all features like Random Forest[20].

In research [21], HOG feature extraction is used to detect objects such as cars. With HOG feature extraction, testing data is created from 304 x 240 pixel images. In research [22], HOG feature extraction with the SVM method is used to classify brain tumors. This method involves three steps in the classification process: pre-processing to resize the image, feature extraction to extract information using HOG feature extraction, and training data using classification testing with SVM, resulting in an accuracy rate of 91%. HOG feature extraction has also been used to identify plant species from leaf patterns using HOG feature extraction and the CNN and KNN method [13], [23], [24].

Previous research has shown that both the Bagging method and HoG feature extraction have yielded good results. In light of this, the current study aims to model the classification of spices by utilizing the Bagging method and HoG feature extraction. The aim is to classify the different types of spices and help the community in identifying them. Such a classification system is expected to benefit young people in learning about different spices.

2. Research Method

2.1 Data Collection Technique

The data for this research was gathered by taking pictures from a Vivo Y35 smartphone camera, while maintaining a distance of 30 cm. To classify the spices, materials had to be purchased from a nearby market and each spice had to be photographed. The chosen model was based on its ability to capture the natural pattern of the spices. The method used to collect spice samples involved placing them in different positions, quantities, and random arrangements on HVS paper.

2.2 Analysis Data

In this study, five different types of spices (Aniseed, Cloves, Cumin, Cardamom, and Candlenuts) were examined, and a total of 750 data points were collected after the data collection process was completed. Each spice type had 150 samples collected for analysis. The data set was divided into two parts for both training and testing purposes, with an 80:20 split ratio. Table 1 shows the accuracy of the data split.

Table 1.
Split data

Types of Spices	The Quantity of Training Samples	The Quantity of Testing Samples
Aniseed	120	30
Cloves	120	30
Cumin	120	30
Cardamon	120	30
Candlenut	120	30

2.3 Architecture Research

Figure 1 illustrates the research methodology adopted in this study

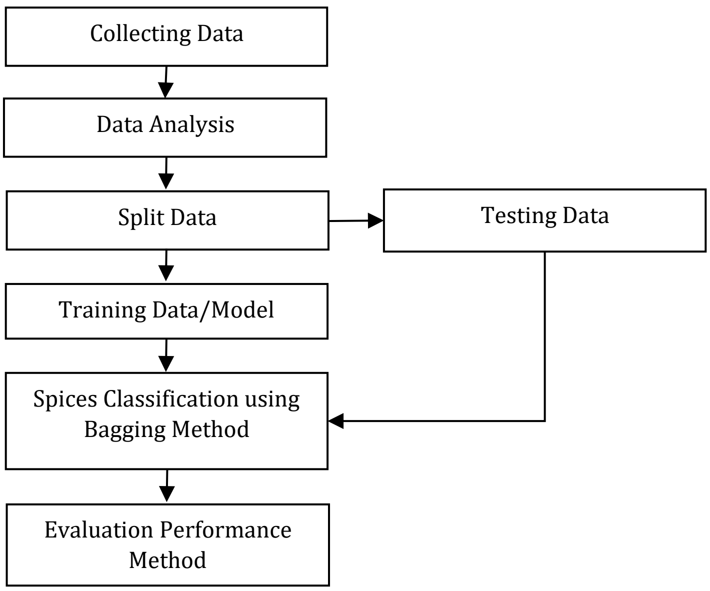


Figure 1. Architecture research[25]

2.4 Evaluation methode

The confusion matrix is a tool used to classify data that has passed or failed a test. It provides information on the actual and predicted classification results, which can be used to evaluate the accuracy of an object estimation model. This study uses various evaluation methods, including Accuracy, Precision, Recall, F1-score, F2-score, and Jaccard score. Accuracy is the simplest measure of performance and is calculated as the ratio of correctly predicted observations to total observations. Precision refers to the proportion of correctly predicted positive observations to the total number of predicted positive observations, whereas recall refers to the proportion of correctly predicted positive observations to all observations in the actual class. The F1-score is a measure that considers false positives and false negatives by calculating a weighted average of precision and recall. The F2-score is a weighted harmonic mean of recall and precision, with a best value of 1 and worst value of 0. Lastly, the Jaccard index, or Jaccard similarity coefficient, measures the similarity between two sets of labels by dividing the size of their intersection by the size of their union[26], [27].

3. Result and Discussion

3.1 Sample of Spices

Aniseed, for example, is a plant with a sweet, licorice-like flavor commonly used in cooking and herbal remedies to aid with digestive problems, coughs, and respiratory ailments. Cloves, on the other hand, are dried flower buds of an evergreen tree, which are widely used in cooking, traditional medicine for their analgesic and anti-inflammatory properties, and in essential oils for aromatherapy and cosmetics. Cumin is a flowering plant with a warm, earthy, and slightly bitter flavor, which is frequently used as a spice in various cuisines, and traditional medicine for its digestive and anti-inflammatory properties. Cardamom, which is a member of the ginger family, has a strong, pungent, and slightly sweet flavor and is used as a spice in many cuisines, as well as in traditional medicine for its digestive and anti-inflammatory properties. Candlenut is a tree that produces seeds used as a food ingredient in numerous Southeast Asian cuisines and in traditional medicine for its anti-inflammatory and analgesic properties. It is important to note that raw Candlenut is toxic and must be roasted or boiled before consumption. Overall, the images in Figure 2 provide a clear visual representation of the various shapes and textures of different spice samples.

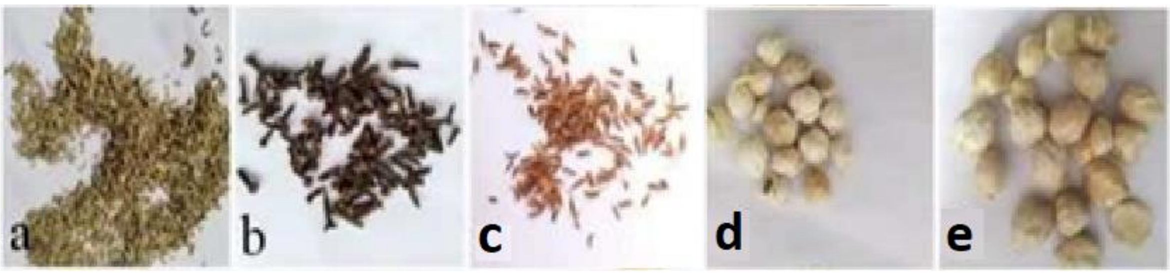


Figure 2. Spice Samples (a) Aniseed, (b) Cloves, (c) Cumin, (d) Cardamom, (e) Candlenuts.

3.2 Histogram of Oriented Gradient Feature Detection

Histogram of Oriented Gradient (HOG) is a feature descriptor used for object detection in computer vision. It was first proposed by Navneet Dalal and Bill Triggs in their 2005 paper "Histograms of Oriented Gradients for Human Detection". The basic idea behind HOG feature detection is to extract local features from an image that capture information about the edges and textures present in the image. This is done by dividing the image into small cells and computing the gradient magnitude and orientation within each cell. The gradient magnitude captures the strength of the edges in the cell, while the orientation captures the direction of the edges.

Once the gradient magnitude and orientation are computed for each cell, a histogram of gradient orientations is created by accumulating the gradient orientations from all the cells within a larger block. The histogram bins are usually chosen to be evenly spaced in the range of 0 to 180 degrees.

The final HOG feature descriptor is created by concatenating the histogram vectors from all the blocks in the image. This results in a high-dimensional feature vector that can be used as input to a machine learning algorithm for object detection.

Overall, HOG is a powerful feature descriptor for object detection that can capture important information about the edges and textures in an image. However, it is computationally intensive and requires careful tuning of parameters such as the cell size, block size, and histogram bin size to achieve optimal performance. The histogram of oriented gradient has generated the detection results, which indicate the identification of objects within an image by analyzing the gradient distribution of pixels. The detection outcomes exhibit a clear identification of the objects within the image, based on the gradient characteristics found by the histogram of oriented gradient method. This technique enables us to more precisely recognize objects within the image, thus facilitating further data processing and analysis.

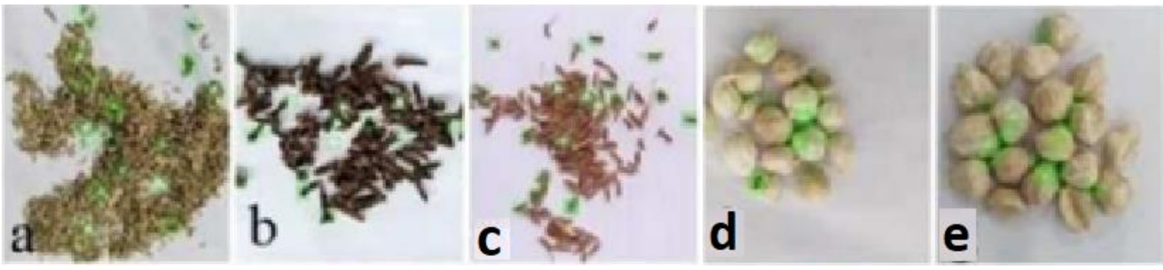


Figure 3. Spice HOG Detection (a) Aniseed, (b) Cloves, (c) Cumin, (d) Cardamom, (e) Candlenuts

3.3 Experiment

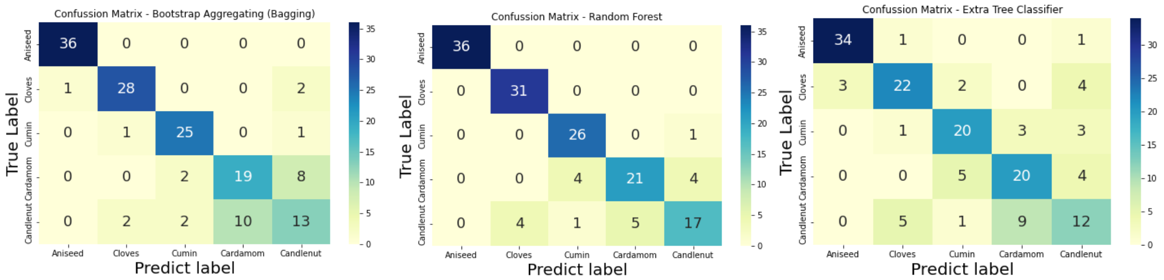
The experiment in this research involved testing a modified version of the bagging method using a specific experimental design shown in Figure 1. Three different models, including Bootstrap Aggregating (Bagging), Random Forests, and Extra Tree Classifier, were utilized to conduct the experiment. The classification outcomes were evaluated and reported through a confusion matrix, which is presented in Figure 4.

This confusion matrix is a representation of the performance of a multiclass classification model that has been trained to classify five different spices: aniseed, cloves, cumin, cardamom, and candlenut. The matrix shows the number of instances of each class that have been correctly or incorrectly classified by the model. The rows of the matrix represent the true classes of the instances, while the columns represent the predicted classes. The diagonal elements of the matrix represent the instances that have been correctly classified, while the off-diagonal elements represent the instances that have been misclassified.

In this confusion matrix (a) Bootstrap Aggregating, we can see that the model has correctly classified most of the instances. For example, all instances of aniseed have been correctly classified as aniseed, and most instances of cloves, cumin, and cardamom have also been correctly classified. However, there are some instances that have been misclassified by the model. For example, two instances of candlenut have been incorrectly classified as cloves, and eight instances of cardamom have been incorrectly classified as candlenut.

Looking at this confusion matrix(b) Random Forests, we can see that the model has correctly classified most of the instances. For example, all instances of aniseed have been correctly classified as aniseed, and most instances of cloves, cumin, and cardamom have also been correctly classified. However, there are still some instances that have been misclassified by the model. For example, four instances of candlenut have been incorrectly classified as cloves, and four instances of cardamom have been incorrectly classified as candlenut. Compared to the previous confusion matrix , this one shows an improvement in the model's performance as there are fewer misclassifications. This could be due to changes made to the model or improvements in the training data.

Looking at this confusion matrix (c) Extra Tree Classifier, we can see that the model's performance is not as good as the previous two confusion matrices. For example, some instances of aniseed have been misclassified as cloves or candlenut, and some instances of cloves have been misclassified as candlenut. There are also instances where multiple misclassifications occurred. For example, 5 instances of candlenut have been misclassified as cloves, and 9 instances of candlenut have been misclassified as cardamom. Overall, this confusion matrix indicates that the model needs further refinement and improvement. The misclassifications could be due to various reasons, such as imbalanced data or insufficient feature representation. Therefore, further data analysis and feature engineering may be needed to improve the model's performance.



(a) (b) (c)

Figure 4. Confusion Matrix of Variations Bagging Algoritmh (a) Bootstrap Aggregating (Bagging), (b) Random Forests, and (c) Extra Tree Classifier

The study did not draw any conclusions based on the classification results obtained from the confusion matrix. Therefore, an evaluation is necessary to assess the degree of success in the classification process.

Table 2.
The performance of Bootstrap Aggregating (Bagging)

	Precision	Recall	F1 - Score	F2 - Score	Jaccard
Aniseed	0.9730	1	0.9863	0.9945	0.9730
Cloves	0.9032	0.9032	0.9032	0.9032	0.8235
Cumin	0.8621	0.9259	0.8929	0.9124	0.8065
Cardamom	0.6552	0.6552	0.6552	0.6552	0.4872
Candlenut	0.5417	0.4815	0.5098	0.4924	0.3421
Accuracy			0.8067		

Precision: The proportion of correctly predicted positive instances out of all predicted positive instances. For instance, the precision of Aniseed is 0.9730, indicating that 97.3% of instances predicted as Aniseed were truly Aniseed.Recall: The proportion of correctly predicted positive instances out of all actual positive instances. For example, the recall of Aniseed is 1, indicating that all instances of Aniseed were correctly predicted.F1-Score: The harmonic mean of precision and recall that balances the two metrics. For example, the F1-Score of Cloves is 0.9032, which is the harmonic mean of its precision and recall.F2-Score: A variant of the F1-Score that emphasizes recall over precision. The F2-Score of Aniseed is 0.9945, indicating that the model has high recall for Aniseed.Jaccard: The similarity between the predicted and actual sets of instances. For example, the Jaccard coefficient of Aniseed is 0.9730, indicating that the predicted set of instances for Aniseed is very similar to the actual set. Accuracy: The proportion of correctly predicted instances out of all instances. The overall accuracy of the model is 0.8067, meaning the model correctly predicted the presence or absence of the spices in 80.7% of instances.

Table 3.
The performance of Random Forests

	Precision	Recall	F1 - Score	F2 - Score	Jaccard
Aniseed	1	1	1	1	1
Cloves	0.8857	1	0.9394	0.9748	0.8857
Cumin	0.8387	0.9630	0.8966	0.9353	0.8125
Cardamom	0.8077	0.7241	0.7636	0.7394	0.6176

	PRECISION	Recall	F1 - Score	F2 - Score	Jaccard
Candlenut	0.7727	0.6296	0.6939	0.6538	0.5313
Accuracy			0.8733		

Precision: Precision measures how accurate the model is in predicting instances that belong to a certain class. A precision score of 1 means that all predictions for that class were correct. In this evaluation, Aniseed has a precision score of 1, which means that all of the instances predicted to be Aniseed were correct. Recall: Recall measures how well the model can identify instances that belong to a certain class. A recall score of 1 means that all instances of that class were correctly identified. In this evaluation, Aniseed has a recall score of 1, which means that all instances of Aniseed were correctly identified. F1-Score: F1-score is a harmonic mean of precision and recall, where a score of 1 indicates perfect precision and recall. In this evaluation, Aniseed has a perfect F1-score of 1, indicating perfect precision and recall. F2-Score: F2-score is similar to F1-score but places more emphasis on recall than precision. In this evaluation, Aniseed has a perfect F2-score of 1, indicating perfect recall and high precision. Jaccard: Jaccard is a measure of the similarity between the predicted set of instances and the actual set of instances. A Jaccard score of 1 means that the predicted set and actual set are identical. In this evaluation, Aniseed has a perfect Jaccard score of 1, indicating perfect similarity between the predicted and actual sets. Overall, the model performed well, with high precision and recall scores for most classes, and an accuracy of 0.8733, indicating a high overall performance.

Table 4.
The performance of Extra Tree Classifier

	PRECISION	Recall	F1 - Score	F2 - Score	Jaccard
Aniseed	0.9189	0.9444	0.9315	0.9392	0.8718
Cloves	0.7586	0.7097	0.7333	0.7190	0.5789
Cumin	0.7143	0.7407	0.7273	0.7353	0.5714
Cardamom	0.6250	0.6897	0.6557	0.6757	0.4878
Candlenut	0.5000	0.4444	0.4706	0.4545	0.3077
Accuracy			0.72		

Precision refers to the proportion of true positive predictions out of all positive predictions made by the model. Recall refers to the proportion of true positive predictions out of all actual positive instances. F1-Score is a harmonic mean of Precision and Recall. F2-Score is a weighted harmonic mean of Precision and Recall, where Recall is given more weight. Jaccard Index measures the similarity between two sets of data. Looking at the results, we can see that Aniseed has the highest Precision, Recall, F1-Score, F2-Score, and Jaccard Index, indicating that the model performs very well on this class. Cloves and Cumin have relatively high scores, but not as high as Aniseed. Cardamom and Candlenut have the lowest scores, indicating that the model does not perform as well on these classes. The overall accuracy of the model is 0.72, which means that the model correctly predicted 72% of the instances in the dataset.

3.4 Discussion

The table 5 shows the evaluation results of three different bagging methods used in a classification problem. Bagging is an ensemble learning technique that involves training multiple models on different subsets of the training data and combining their predictions to improve the overall performance of the model.

The three bagging methods evaluated in this study are Bootstrap Aggregating (Bagging), Random Forests, and Extra Tree Classifier. The evaluation metrics used in this study include precision, recall, F1-score, F2-score, Jaccard index, and accuracy. From the evaluation results, it can be seen that Random Forests outperformed the other two bagging methods in most of the evaluation metrics. Random Forests achieved the highest precision, recall, F1-score, F2-score, Jaccard index, and accuracy values of 0.861, 0.8633, 0.8587, 0.8607, 0.7694 respectively, indicating that it is the most effective bagging method for the spicess classification problem in this study.

Bootstrap Aggregating (Bagging) and Extra Tree Classifier achieved lower performance compared to Random Forests in most of the evaluation metrics. However, they still achieved decent performance, with Bootstrap Aggregating (Bagging) achieving an accuracy of 0.8067 and Extra Tree Classifier achieving an accuracy of 0.72.

Overall, the classification results obtained from the confusion matrix suggest a decent level of success in classifying different types of spices using HOG feature extraction and bagging method. However, further evaluation of model performance is necessary, such as increasing the amount of training data or considering alternative feature extraction methods to improve the accuracy of classification outcomes.

Table 5.
The performance of Variations Bagging Method

Bagging Method	Precision	Recall	F1 - Score	F2 - Score	Jaccard	*Accuracy*
Bootstrap Aggregating (Bagging)	0.7870	0.7932	0.7895	0.7915	0.6864	0.8067
Random Forests	**0.8610**	**0.8633**	**0.8587**	**0.8607**	**0.7694**	**0.8733**
Extra Tree Classifier	0.7034	0.7058	0.7037	0.7047	0.5635	0.72

Table 6.
Comparison with previous research

Bagging Method	Precision	Recall	F1 - Score	F2 - Score	Jaccard	*Accuracy*
Bootstrap Aggregating (Bagging)	0.7870	0.7932	0.7895	0.7915	0.6864	0.8067
Random Forests	**0.8610**	**0.8633**	**0.8587**	**0.8607**	**0.7694**	**0.8733**
Extra Tree Classifier	0.7034	0.7058	0.7037	0.7047	0.5635	0.72
Adaboost Classifier [28]	0.7094	0.6895	0.6821	0.6821	0.5338	0.6800
Gradient Boosting Classifier [28]	0.7538	0.7568	0.7537	0.7552	0.6289	0.7600
XGB Classifier [28]	0.8117	0.8088	0.8095	0.8089	0.6951	0.8060
Ligth GBM Classifier[28]	0.7782	0.7763	0.7733	0.7742	0.6440	0.7730

4. Conclusion

The study found that all four tested models had relatively high accuracy in classifying different types of spices, with the Random Forest model performing the best, achieving an accuracy of 87.33%. The Random Forest model also outperformed the other models in terms of precision, recall, F1-score, and F2-score. On the other hand, the Extra Tree Classifier had the lowest performance, with precision, recall, F1-score, and F2-score values of 0.7034, 0.7058, 0.7037, and 0.7047, respectively. Furthermore, the Jaccard Score, which evaluates the accuracy of clustering algorithms by measuring the similarity between two sets of data, ranged from 0.5635 to 0.7694 in this study, indicating some overlap between predicted and actual class labels. Based on the evaluation metrics presented, Random Forests outperformed the other two methods with the highest precision, recall, F1-score, F2-score, Jaccard, and accuracy values. This indicates that Random Forests has a better performance in predicting the target variable compared to Bootstrap Aggregating (Bagging) and Extra Tree Classifier. However, it is important to note that the difference in performance between these methods may be dependent on the specific dataset and problem being solved. Therefore, further analysis and experimentation may be necessary to determine the best method for a given scenario. Overall, the study suggests that the HOG feature extraction method combined with Boosting algorithms can effectively classify different types of spices. However, improving the sample size and considering alternative feature extraction methods may enhance the accuracy of the classification outcomes. Further evaluations may be necessary to assess the generalizability of these models to other datasets or to determine if retraining is required.

References

[1] H. Anggrasari and W. A. Saputro, "COMPARATIVE ADVANTAGE OF INDONESIA WITH COMPETITIVE COUNTRIES FOR EXPORTING OF WORLD SPICES," *Journal ASEAN Dynamics and Beyond*, vol. 2, no. 1, 2021.

[2] S. Y. Riska and L. Farokhah, "Klasifikasi Bumbu Dapur Indonesia Menggunakan Metode K-Nearest Neighbors (K-NN)," *SMATIKA*, vol. 11, no. 1, 2021.

[3] M. E. Embuscado, "Spices and herbs: Natural sources of antioxidants – a mini review," *J Funct Foods*, vol. 18, pp. 811–819, 2015, doi: https://doi.org/10.1016/j.jff.2015.03.005.

[4] A. Ali, H. Wu, E. N. Ponnampalam, J. J. Cottrell, F. R. Dunshea, and H. A. R. Suleria, "Comprehensive Profiling of Most Widely Used Spices for Their Phenolic Compounds through LC-ESI-QTOF-MS2 and Their Antioxidant Potential," *Antioxidants*, vol. 10, no. 5, 2021, doi: 10.3390/antiox10050721.

[5] T. Sudhakar Johnson and D. B. A. Narayana, "Role of spices in offering natural immunity to fight various diseases," *Pharmacognosy Journal*, vol. 13, no. 2. EManuscript Technologies, pp. 600–613, Mar. 01, 2021. doi: 10.5530/pj.2021.13.76.

[6] K. Umakanth and S. A. Girija, "ROLE OF COMMONLY USED ASIAN SPICES IN BOOSTING IMMUNITY AGAINST INFECTIOUS AGENTS," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 1, 2020.

[7] M. Setiawan, M. Rahayu, D. Q. Wahyu Ningsih, and N. Arifa, "Tumbuhan Rempah dan Masakan Tradisional di Kelurahan Nanggewer Mekar, Cibinong, Kabupaten Bogor," *Jurnal Masyarakat dan Budaya*, vol. 23, no. 3, Mar. 2022, doi: 10.14203/jmb.v23i3.1434.

[8] I. Wulandari, H. Yasin, and T. Widiharih, "KLASIFIKASI CITRA DIGITAL BUMBU DAN REMPAH DENGAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN)," *JURNAL GAUSSIAN*, vol. 9, no. 3, 2020, [Online]. Available: https://ejournal3.undip.ac.id/index.php/gaussian/

[9] M. Muhathir, M. H. Santoso, and R. Muliono, "Analysis Naïve Bayes In Classifying Fruit by Utilizing Hog Feature Extraction," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 4, no. 1, pp. 151–160, Jul. 2020, doi: 10.31289/jite.v4i1.3860.

[10] N. Dewi Girsang and M. Muhathir, "Classification Of Batik Images Using Multilayer Perceptron With Histogram Of Oriented Gradient Feature Extraction," in *PROC. INTERNAT. CONF. SCI. ENGIN*, 2021, vol. 4.

[11] S. Ghaffari, P. Soleimani, K. F. Li, and D. W. Capson, "Analysis and Comparison of FPGA-Based Histogram of Oriented Gradients Implementations," *IEEE Access*, vol. 8, pp. 79920–79934, 2020, doi: 10.1109/ACCESS.2020.2989267.

[12] J. P. Tanjung and M. Muhathir, "Classification of facial expressions using SVM and HOG," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 3, no. 2, pp. 210–215, Jan. 2020, doi: 10.31289/jite.v3i2.3182.

[13] T. Tri Saputra Sibarani and C. Author, "Analysis K-Nearest Neighbors (KNN) in Identifying Tuberculosis Disease (Tb) By Utilizing Hog Feature Extraction," *International of Computer Science and Information Technology (AIoCSIT) JournalISSN*, vol. 1, no. 1, pp. 33–38, 2020.

[14] S. Bakheet and A. Al-Hamadi, "A Framework for Instantaneous Driver Drowsiness Detection Based on Improved HOG Features and Naïve Bayesian Classification," *Brain Sci*, vol. 11, no. 2, 2021, doi: 10.3390/brainsci11020240.

[15] P. Xu, Z. Xiao, X. Wang, L. Chen, C. Wang, and F. An, "A multi-core object detection coprocessor for multi-scale/type classification applicable to iot devices," *Sensors (Switzerland)*, vol. 20, no. 21, pp. 1–13, Sep. 2020, doi: 10.3390/s20216239.

[16] Muhathir, R. A. Rizal, J. S. Sihotang, and R. Gultom, "Comparison of SURF and HOG extraction in classifying the blood image of malaria parasites using SVM," in *2019 International Conference of Computer Science and Information Technology (ICoSNIKOM)*, 2019, pp. 1–6. doi: 10.1109/ICoSNIKOM48755.2019.9111647.

[17] H. Leidiyana and J. Warta, "Implementasi Metode SVM untuk Klasifikasi Bunga dengan Ekstraksi Fitur Histogram of Gradient (HOG)," *Journal of Informatic and Information Security*, vol. 3, no. 1, pp. 89–98, 2022.

[18] L. Breiman, "Bagging predictors," *Mach Learn*, vol. 24, no. 2, pp. 123–140, 1996, doi: 10.1007/BF00058655.

[19] L. Breiman, "Random Forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[20] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach Learn*, vol. 63, no. 1, pp. 3–42, Apr. 2006, doi: 10.1007/s10994-006-6226-1.

[21] F. A. I. Achyunda Putra, F. Utaminingrum, and W. F. Mahmudy, "HOG Feature Extraction and KNN Classification for Detecting Vehicle in The Highway," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 14, no. 3, p. 231, Jul. 2020, doi: 10.22146/ijccs.54050.

[22] T. Adilah and Q. N. Azizah, "Klasifikasi Tumor Otak Menggunakan Ekstraksi Fitur HOG dan Support Vector Machine," 2022. [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/infortech45

[23] T. Q. Bao, N. T. T. Kiet, T. Q. Dinh, and H. X. Hiep, "Plant species identification from leaf patterns using histogram of oriented gradients feature space and convolution neural networks," *Journal of Information and Telecommunication*, vol. 4, no. 2, pp. 140–150, 2020, doi: 10.1080/24751839.2019.1666625.

[24] N. Khairina, T. T. S. Sibarani, R. Muliono, Z. Sembiring, and M. Muhathir, "Identification of Pneumonia using The K-Nearest Neighbors Method using HOG Fitur Feature Extraction," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 5, no. 2, pp. 562–568, Jan. 2022, doi: 10.31289/jite.v5i2.6216.

[25] F. Fadlisyah and M. Muhathir, "Performance Evaluation Of Variations Boosting Algorithms For Classifying Formalin Fish From Photos," *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 6, no. 2, pp. 621–629, Jan. 2023, doi: 10.31289/jite.v6i2.6614.

[26] Muhathir and Al-Khowarizmi, "Measuring the Accuracy of SVM with Varying Kernel Function for Classification of Indonesian Wayang on Images," in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 1190–1196. doi: 10.1109/DASA51403.2020.9317197.

[27] M. Ula, M. Muhathir, and I. Sahputra, "Optimization of Multilayer Perceptron Hyperparameter in Classifying Pneumonia Disease Through X-Ray Images with Speeded-Up Robust Features Extraction Method," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 13, no. 10, 2022, [Online]. Available: www.ijacsa.thesai.org

[28] Muhathir, M., Trisno Pangestu, R., Safira, I., & Melisah, M. (2023). Performance Comparison of Boosting Algorithms in Spices Classification Using Histogram of Oriented Gradient Feature Extraction. *Journal of Computer Science, Information Technology and Telecommunication Engineering (JCoSITTE)*, *4*(1), 342–349. https://doi.org/10.30596/jcositte.v4i1.13710