Research Article

Wheat Seed Classification: Utilizing Ensemble Machine Learning Approach

Ajay Khatri ¹, Shweta Agrawal,² and Jyotir M. Chatterjee ³

¹CSE, SAGE University, Indore 452001, India

²CSE, SAGE University, Indore 452001, India

³Lord Buddha Education Foundation, Kathmandu 44600, Nepal

Correspondence should be addressed to Jyotir M. Chatterjee; jyotirchatterjee@gmail.com

Received 19 October 2021; Accepted 13 January 2022; Published 2 February 2022

Academic Editor: Punit Gupta

Copyright © 2022 Ajay Khatri et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recognizing and authenticating wheat varieties is critical for quality evaluation in the grain supply chain, particularly for methods for seed inspection. Recognition and verification of grains are carried out manually through direct visual examination. Automatic categorization techniques based on machine learning and computer vision offered fast and high-throughput solutions. Even yet, categorization remains a complicated process at the varietal level. The paper utilized machine learning approaches for classifying wheat seeds. The seed classification is performed based on 7 physical features: area of wheat, perimeter of wheat, compactness, length of the kernel, width of the kernel, asymmetry coefficient, and kernel groove length. The dataset is collected from the UCI library and has 210 occurrences of wheat kernels. The dataset contains kernels from three wheat varieties Kama, Rosa, and Canadian, with 70 components chosen at random for the experiment. In the first phase, K-nearest neighbor, classification and regression tree, and Gaussian Naïve Bayes algorithms are implemented for classification. The results of these algorithms are compared with the ensemble approach of machine learning. The results reveal that accuracies calculated for KNN, decision, and Naïve Bayes classifiers are 92%, 94%, and 92%, respectively. The highest accuracy of 95% is achieved through the ensemble classifier in which decision is made based on hard voting.

1. Introduction

In many developing nations, farming is the most significant economic sector. Most of the tasks are carried out without the use of modern technology. Seed categorization is usually done based on human understanding. The purification of seeds plays an important role in this process and must be enhanced. Manually determining the type of wheat needs expert judgment and takes time. When an array of seeds appears so similar, manually distinguishing them becomes a challenging process [1–4].

A quality evaluation method for wheat crops is required with the evolution of the grain chain [5]. The goal is a high-quality wheat product in greater quantities. Tests on seed germination are required for seed labeling [6]. The seeds were tested for pureness test, which is required to ascertain a seed sample's physical and varietal purity. The genetic

integrity of the original wheat cultivar may be compromised by mechanical mixing and improper labeling [7].

Classification testing is accomplished via a taxonomic categorization method and nondestructive grain feature analysis [8, 9]. Seed tester classifies seeds on two levels: species level and varietal level for varietal purity. The varietal level may be challenging due to the great degree of resemblance in the characteristics of different kinds of wheat seeds. The growth circumstances may also affect the grain's properties [5]. The actual classification test procedure is still a low-throughput process, and its correctness is contingent on the expert's performance and cumulative experience.

Machine learning techniques are now the subject of study in a variety of disciplines, particularly with the expansion of the Internet and the usage of bigger datasets [4]. Without specific tools or automatic software procedures, it is difficult for a human operator to interpret or handle such

data. Machine learning is frequently used in various applications such as categorization, regression, and forecasting to meet these demands [10]. The use of a single classifier for the objects which has a very minute difference in physical characteristics such as color, texture, and morphology does not give better accuracy [11]. To address this problem, ensemble approach is used in the present work. By combining many models into a single, highly reliable model, ensemble methods seek to increase model prediction. The most common ensemble approaches are boosting, bagging, and stacking. Ensemble techniques are particularly well suited to regression and classification, where they minimize bias and variance while increasing model accuracy. The major findings of this paper are as follows:

- (1) An optimized classifier is designed for wheat seed classification by utilizing an ensemble machine learning approach with bagging
- (2) The model is compared with three machine learning classifiers: K-nearest neighbors (KNN) classifier, decision tree classifier (CART), and Gaussian NB (NB).
- (3) The highest achieved accuracy is 95%, which is with the ensemble method

2. Related Work

Seed categorization using machine algorithms has been the subject of research. These studies employed a variety of machine classifiers and achieved a high degree of accuracy in their work. Machine learning techniques had previously been successfully applied in a variety of production chains for seed and cereals classification [12–14]. In [15], the study shows the capability and possibilities of machine vision for shapes, sizes, and varietal types using well-trained multilayer neural network classifiers. They utilized Weka classification tools such as function, Bayes, meta, and lazy approaches to categorize the seeds. In [16], the authors proposed a fuzzy theory-based approach for recognizing wheat seed types that take into account the features of the seed. The tabu search technique was used. In [17], the authors have used an artificial neural network for classifying wheat seeds based on VLC and obtained an accuracy of 92.1 percent and 85.72 percent, respectively. In [18], authors have discussed morphological, color, and textural characteristics of the seed. If there is a very minute difference in morphological features, then seed classification is very difficult. Cereal yield is determined by the number of grains per ear and the size of the grains. Counting seeds and morphometry by sight is time-consuming. As a result, different ways for effective grain morphometry employing image processing techniques have been proposed [19, 20].

In [21], the authors created a workstation to aid in grain analysis for classification, and a video colorimetry approach is presented to support in determining cereal grain color. The categorization of chickpea seed varieties was done based on the morphological characteristics of chickpea seeds, using 400 samples from four types: Kaka, Piroz, Ilc, and Jam [22]. According to the commercial point of view, a machine vision

built of existing neural network models may be utilized for rice quality assessment [23]. In this, it uses neural networks to categorize rice varieties, using a total of nine separate varieties of rice. The authors employ seed image acquisition to classify these variations. They also created a method for extracting 13 morphological features, 6 color features, and 15 texture features from color photographs of seeds. Their model has produced an overall classification accuracy of 92%. The k-nearest neighbors classifier necessitates storing the entire training set, which can be prohibitively expensive when the set is huge, and several researchers have attempted to eliminate the training set's redundancy to relieve this problem [24, 25].

For plant categorization, the authors have utilized deep learning models [26]. Two tendencies may be seen in the current state of the art. The first is linked to high-throughput phenotyping and plant identification, as evidenced by Ubbens and Stavness' work in this area [27]. The second problem is plant disease identification and monitoring [28, 29]. In [30], the authors present many voting techniques for testing ensembles of classifiers learned using the bagging approach. Multilayer perceptron is used as classifiers. Using groups of classifiers rather than individual ones is one option. Bagging [31] and boosting [32] are two of the most well-known ensemble techniques, in which many classifiers are combined to generate a single, more accurate result. In [33], the authors studied the performance of several voting techniques, with bagging being utilized for the reconciliation model, which is a process of merging classification models. Table 1 contains various features considered in machine vision systems for food grain quality evaluation.

3. Methodology

The methodology adopted for this work includes the collection of datasets, features identification, data augmentation, classification using machine learning algorithms like KNN, Naive Bayes, and CART implementation, implementation of ensemble approach for better accuracy with fine-tuning, and comparison of results.

3.1. Dataset. In the study, the seed dataset was received from the UCI library [42]. There are 210 occurrences of wheat kernels in the collection. In addition to the class attribute, each instance contains 7 other attributes. All the samples share the same 7 characteristics (area, perimeter, compactness, kernel length, kernel width, asymmetry coefficient, and kernel groove length). All of the characteristics are constant. Figure 1 shows the features of the dataset number of classes presented and selected machine learning algorithms.

The examined set contained kernels from three wheat varieties: Kama, Rosa, and Canadian, with 70 components. To perceive high-quality picturing of the interior kernel construction, a soft X-ray method is employed. It is less abrasive and more affordable than other sophisticated imaging techniques such as scanning microscopy or laser technology. Images are recorded on X-ray KODAK plates

TABLE 1: Features considered in machine vision systems for food grain quality evaluation.

S. no.	Features	Application	Grain	References
1	Morphology	Variety-based classification	Rice	[34–36]
2	Color	Identification of microbial infection	Popcorn	[37]
3	Morphology + color	Variety-based classification	Chickpea	[38, 39]
4	Morphology	Variety-based classification	Wheat	[40, 41]

FIGURE 1: Count of different wheat seed varieties.

measuring 13×18 cm. At the Institute of Agrophysics of the Polish Academy of Sciences in Lublin, studies were supplemented by association harvested wheat grain patenting from experimental fields. In Table 2, classes represent 3 varieties of wheat which are Kama (1), Rosa (2), and Canadian (3).

3.2. Machine Learning Models

K-nearest neighbor: a k-nearest-neighbor algorithm is a data categorization method that predicts how likely a data point belong to one of two groups based on which group the data points closest to it belong to [43]. Any common method can be used to determine the distance. A Euclidean distance is an example of a distance. After that, we collect a certain feature value from all the training set in the immediate vicinity. We categorize our fresh testing data using most of this number as a prediction [44].

Classification and regression trees (CART): CART is a nonparametric supervised learning method [45, 46]. The objective is to build a model of the value of a target variable by learning basic decision rules from data characteristics. A piecewise constant is approximated by a tree. Nonstatisticians can analyze CART quite well [45, 47].

Gaussian Naïve Bayes: A basic approach for Naive Bayes is Bayes decision theory. The likelihood is used to make this classification decision. The posterior probability is calculated using the likelihood, prior probabilities, and evidence. Evidence is merely a scalar that ensures that posterior probability equals one. Resultant

classes for the given test data are chosen based on the category with the highest posterior probability.

Ensemble methods: ensemble methods are strategies for increasing the accuracy of model outputs by integrating many models rather than utilizing just one. The integrated models greatly improve the accuracy of the outcomes. The popularity of ensemble techniques in machine learning has risen because of this. When dealing with enormous amounts of data or a lack of appropriate data, ensemble-based solutions can be unexpectedly effective. When the quantity of training data is too enormous to train a single classifier, the data can be divided into smaller groups deliberately. After that, each division may be used to train a distinct classifier, which can then be merged using a suitable combination algorithm. If there is not enough data, bootstrapping may be used to train alternative classifiers using distinct bootstrap samples of the data, each of which is a random sample of the data taken with replacement and handled as if it were drawn independently from the underlying distribution [48].

Bagging is one of the oldest, more obvious, and most likely simplest ensemble-based methods, and it has very maximum performance. Bootstrapped copies of the training data are used to obtain the diversity of classifiers in bagging such that different chunks of training data are arbitrarily chosen from the entire training dataset, with substitution. Each subset of training data is utilized to train a particular sort of classifier. Individual classifiers are then merged using a simple majority vote. By resampling the data, boosting also produces an ensemble of classifiers, which are subsequently

TABLE 2: Sample from wheat seed dataset.

Area of wheat	Perimeter of wheat	Compactness	Kernel length	Kernel width	Asymmetry coefficient	Kernel groove length	Classes
18.55	16.22	0.8865	6.153	3.674	1.738	5.894	2
19.15	16.45	0.889	6.245	3.815	3.084	6.185	2
14.11	14.1	0.8911	5.42	3.302	2.7	5	1
20.88	17.05	0.9031	6.45	4.032	5.016	6.321	2
11.84	13.21	0.8521	5.175	2.836	3.598	5.044	3
11.87	13.02	0.8795	5.132	2.953	3.597	5.132	3
15.78	14.91	0.8923	5.674	3.434	5.593	5.136	1

merged by majority vote. Resampling, on the other hand, is used deliberately in boosting to give the most useful training data for each subsequent classifier.

We applied a hard voting classifier, which means that the forecasted output class is the one that obtains the most votes. Three classifiers predicted the output wheat classes Kama, Rosa, and Canadian, and most of them anticipated Kama wheat variety as the result. As a result, the ultimate forecast will be Kama wheat.

3.3. Experimental Setup. Python3 libraries such as NumPy, SciPy, scikit-learn, Keras, pandas, and Matplotlib are utilized to perform the categorization through ML models. Scikit-learn appears to be the most user-friendly and reliable machine learning library [49, 50]. The foundations of this package are NumPy, SciPy, and Matplotlib. The results of the dataset analysis, as well as the model's training and testing using numerous feature

extractions are presented in this section. Figure 2 shows the classification process.

3.4. Performance Evaluation Matrix. To measure the effectiveness of the classifier, a confusion matrix is required, which provides the number of accurate and wrong predictions based on known true values. It shows true positive (TP): actual value is true and model predicted it true, true negative (TN): both actual value and predicted value are false, false positive (FP): actual value is false but model predicted it true, and false negative (FN): actual value is true but model predicted it as false.

Accuracy: accuracy is the measure of how often a model has predicted the right value as per the given input. But it does not give detailed information regarding FP and FN. For some applications where FP and FN are considerable, *F1*-score and recall play a very important role. Accuracy is calculated by

$$\text{accuracy} = \frac{\text{true positive} + \text{true negative}}{\text{true positive} + \text{true negative} + \text{false positive} + \text{false negative}}. \quad (1)$$

Precision: this evaluation parameter tells how frequently a model predicts true positive. The low value of precision infers high false positives. Formula for calculating precision is as follows:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (2)$$

Recall: this parameter gives information regarding how often a model predicts false negatives. The low value of recall means the model predicted high false negatives. Formula for calculating recall is as follows:

$$\text{recall} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN}}. \quad (3)$$

F1-score: the *F1*-score is calculated by combining both precision and recall. That is, a high *F1* score indicates a low number of false positives and false negatives, which infers that the model is accurately detecting actual threats and are not bothered by false alarms. The formula for calculating the *F1* score is

$$\text{F1 - score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \quad (4)$$

4. Result

Results present the evaluation of considered four ML models: KNN, Naive Bayes, CART, and ensemble method. For evaluating models, we have divided the collected data into 70% for the training set and 30% for the testing set. The wheat classes Kama, Rosa, and Canadian have been assigned numbers 1, 2, and 3, correspondingly. The results of all considered algorithms are evaluated based on recall, precision, *F1*-score, and accuracy. Tables 3–7 show the value of these parameters for KNN, Naive Bayes, CART, and ensemble method, respectively. Figures 3–6 show the confusion matrix for KNN, Naive Bayes, CART, and ensemble method, respectively. The accuracy determined for KNN, decision, and Naive Bayes classifiers is 92 percent, 94 percent, and 92 percent, respectively, according to the data. Ensemble classifier, which makes decisions based on hard voting, has the best accuracy of 95 percent.

FIGURE 2: The block diagram of the classification process.

TABLE 3: Result of wheat seed classification using KNN.

Class	Precision	Recall	F1-score	Support
1	0.9	0.86	0.88	22
2	0.92	0.96	0.94	23
3	0.94	0.94	0.94	18
Avg/total	0.92	0.92	0.92	63

TABLE 4: Result of wheat seed classification using CART.

Class	Precision	Recall	F1-score	Support
1	0.91	0.91	0.91	22
2	0.96	0.96	0.96	23
3	0.94	0.94	0.94	18
Avg/total	0.94	0.94	0.94	63

TABLE 5: Result of wheat seed classification using NB.

Class	Precision	Recall	F1-score	Support
1	0.87	0.91	0.89	22
2	1	0.91	0.95	23
3	0.89	0.94	0.92	18
Avg/total	0.92	0.92	0.92	63

TABLE 6: Hard voting in ensemble method.

S. no.	KNN	CART	Naive Bayes (NB)	Ensemble method
1	1	3	1	1
2	3	2	3	3
3	2	2	3	2
4	3	3	1	3

TABLE 7: Result of wheat seed classification using ensemble method.

Class	Precision	Recall	F1-score	Support
1	0.91	0.95	0.93	22
2	1	0.96	0.98	23
3	0.94	0.94	0.94	18
Avg/total	0.95	0.95	0.95	63

5. Discussion

To implement the KNN, we use the scikit-learn K-neighbors classifier. As an input parameter, the approach requires the number of neighbors. We can

determine the wheat seed category by simply changing the number of neighbors. With this procedure, 92% accuracy is attained. The accuracy of building a CART using the scikit-learn is 94 percent. We used the scikit-learn Gaussian NB classifier and 92% accuracy is attained. We

FIGURE 3: Confusion matrix (KNN).

FIGURE 4: Confusion matrix (CART).

FIGURE 5: Confusion matrix (NB).

have achieved 95% accuracy using the ensemble method. Based on machine learning, a model for predicting determining factors of seeds has been created. The precision,

recall, and F1 score of wheat seed dataset categorization using K-neighbors classifier (KNN), classification and regression trees (CART), Gaussian NB (NB), and

FIGURE 6: Confusion matrix (ensemble method).

TABLE 8: Comparison between precision, recall, and F1-score

Classifier	Precision	Recall	F1-score
KNN	0.92	0.92	0.92
CART	0.94	0.94	0.94
NB	0.92	0.92	0.92
Ensemble	0.95	0.95	0.95

TABLE 9: Summary of various classifiers accuracy.

Classifier	Accuracy (%)
KNN	92
CART	94
NB	92
Ensemble	95

FIGURE 7: Summary of various classifier accuracy.

ensemble method are compared in Table 8. Table 9 shows the summary of various classifiers accuracy. A chart is also presented to compare the accuracy of various methods in Figure 7 which shows the summary of various classifier accuracy.

6. Conclusion

Machine learning approaches in grain seed analysis and classification are playing a very important role. The major challenge in seed classification is the very minute difference

between different categories of seeds. The accuracy of predictions with this challenge is improved by utilizing the concepts of ensemble learning. Wheat seed classification by considering seven independent features area, perimeter, compactness, kernel length, kernel width, asymmetry coefficient, and kernel groove length is presented in the paper. The ensemble machine learning approach with bagging and hard voting is utilized to best fit the classifier. Three machine learning algorithms K-nearest neighbors classifier (KNN), classification and regression trees (CART), and Gaussian NB (NB) are also implemented to compare the results. The

results reveal that the accuracy calculated for KNN, decision, and Naïve Bayes classifiers are 92%, 94%, and 92%, respectively. The highest accuracy 95% is achieved through the ensemble classifier in which the decision is made based on hard voting. In the future, we can use other classification algorithms to improve accuracy.

Data Availability

Datasets related to this article can be found at “<https://archive.ics.uci.edu/ml/datasets/seeds>,” an open-source online data repository hosted at UCI Machine Learning Repository [42].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] R. Tandon, D. S. Agrawal, and D. P. Goyal, “Sequential convolutional neural network for automatic breast cancer image classification using histopathological images,” *J. Crit. Rev.* vol. 7, no. 15, p. 14, 2020.
- [2] H. S. H. Chitra, S. Suguna, and S. N. Sujatha, “A survey on image analysis techniques in agricultural product,” *Indian Journal of Science and Technology*, vol. 9, no. 12, 2016.
- [3] M. R. Siddagangappa and A. P. A. H. Kulkarni, “Classification and quality analysis of food grains,” *IOSR Journal of Computer Engineering*, vol. 16, no. 4, pp. 1–10, 2014.
- [4] S. Agrawal and S. K. Jain, “Medical text and image processing: applications, issues and challenges,” in *Machine Learning with Health Care Perspective*, V. Jain and J. M. Chatterjee, Eds., vol. 13, pp. 237–262, Cham: Springer International Publishing, Manhattan, NY, USA, 2020.
- [5] C. Wrigley, “Cereal-grain morphology and composition,” in *Cereal Grains*, C. Wrigley, I. Batey, and D. Miskelly, Eds., pp. 55–87, Woodhead Publishing, Cambridge, UK, 2nd edition, 2017.
- [6] S. G. Elias, L. O. Copeland, M. B. McDonald, and R. Z. Baalbaki, *Seed Testing: Principles and Practices*, East Lansing: Michigan State University Press, East Lansing, MI, USA, 2012.
- [7] S. Matthews, “Copeland, L.O. and McDonald, M.B. Principles of seed science and technology. 4th edn,” *Annals of Botany*, vol. 89, no. 6, p. 798, 2002.
- [8] W. S. Meyer and H. D. Barrs, “Non-destructive measurement of wheat roots in large undisturbed and repacked clay soil cores,” *Plant and Soil*, vol. 85, no. 2, pp. 237–247, 1985.
- [9] J. Acer, *Rules Proposals for the International Rules for Seed Testing*, Vol. 47, International Seed Testing Association, Switzerland, Europe, 2019 edition, 2019.
- [10] M. Agrawal and S. Agrawal, “A systematic review on artificial intelligence/deep learning applications and challenges to battle against COVID-19 pandemic,” *Disaster Advances*, vol. 14, no. 8, pp. 90–99, 2021.
- [11] N. S. Visen, J. Paliwal, D. S. Jayas, and N. D. G. White, “AE-automation and emerging technologies,” *Biosystems Engineering*, vol. 82, no. 2, pp. 151–159, 2002.
- [12] D. I. Patrício and R. Rieder, “Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review,” *Computers and Electronics in Agriculture*, vol. 153, pp. 69–81, 2018.
- [13] P. Vithu and J. A. Moses, “Machine vision system for food grain quality evaluation: a review,” *Trends in Food Science & Technology*, vol. 56, pp. 13–20, 2016.
- [14] C.-J. Du and D.-W. Sun, “Learning techniques used in computer vision for food quality evaluation: a review,” *Journal of Food Engineering*, vol. 72, no. 1, pp. 39–55, 2006.
- [15] T. Tujo, “A predictive model to predict seed classes using machine learning,” *Int. J. Eng. Tech. Res.*, vol. 6, pp. 334–344, 2019.
- [16] L. Li and S. Liu, “Wheat cultivar classifications based on tabu search and fuzzy C-means clustering algorithm,” in *Proceedings of the 2012 4th International Conference on Computational and Information Sciences*, pp. 493–496, Chongqing, China, August. 2012.
- [17] R. Choudhary, S. Mahesh, J. Paliwal, and D. S. Jayas, “Identification of wheat classes using wavelet features from near infrared hyperspectral images of bulk samples,” *Biosystems Engineering*, vol. 102, no. 2, pp. 115–127, 2009.
- [18] S. V. Neeraj Singh Visen, J. Jitendra Paliwal, D. Digvir Jayas, and N. D. G. White, “Image analysis of bulk grain samples using neural networks,” *Canadian Biosystems Engineering/Le Génie des biosystems au Canada*, vol. 46, 2003.
- [19] T. Tanabata, T. Shibaya, K. Hori, K. Ebana, and M. Yano, “SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis,” *Plant Physiology*, vol. 160, no. 4, pp. 1871–1880, 2012.
- [20] E. Komyshev, M. Genaev, and D. Afonnikov, “Evaluation of the SeedCounter, A mobile application for grain phenotyping,” *Frontiers of Plant Science*, vol. 7, 2017.
- [21] M. R. Neuman, E. Shweddyk, and W. Bushuk, “A PC-based colour image processing system for wheat grain grading,” in *Proceedings of the 3rd International Conference on Image Processing and its Applications*, 1989, pp. 242–246, Warwick, UK, July 1989.
- [22] S. Salah Ghamari, “Classification of chickpea seeds using supervised and unsupervised artificial neural networks,” *African Journal of Agricultural Research*, vol. 7, no. 21, 2012.
- [23] C. Silva and U. Sonnadara, “Classification of rice grains using neural networks,” *Proceedings of Technical Sessions*, vol. 29, 2013.
- [24] M. Kubat and J. R. Martin Cooperson, “Voting Nearest-Neighbor Subclassifiers,” in *Proceedings of the 17th International Conference on Machine Learning*, pp. 503–510, San Francisco, CA, USA, January. 2000.
- [25] E. Alpaydin, “Voting over Multiple Condensed Nearest Neighbors,” in *Lazy Learning*, D. W. Aha, Ed., Springer Netherlands, Dordrecht, The Netherlands, 1997.
- [26] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: a survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [27] J. R. Ubbens and I. Stavness, “Corrigendum: deep plant phenomics: a deep learning platform for complex plant phenotyping tasks,” *Frontiers of Plant Science*, vol. 8, p. 2245, 2018.
- [28] M. Brahimi, K. Boukhalfa, and A. Moussaoui, “Deep learning for tomato diseases: classification and symptoms visualization,” *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, 2017.
- [29] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
- [30] M. Van Erp, L. Vuurpijl, and L. Schomaker, “An overview and comparison of voting methods for pattern recognition,” in *Proceedings of the 8th International Workshop on Frontiers in*

- Handwriting Recognition, Niagara on the Lake*, pp. 195–200, Niagara-on-the-Lake, ON, Canada, August 2002.
- [31] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
 - [32] R. E. Schapire and Y. Singer, “Improved boosting algorithms using confidence-rated predictions,” in *Proceedings of the 11th annual conference on Computational learning theory - COLT’98*, pp. 80–91, Madison, WI, USA, July 1998.
 - [33] K. T. Leung and D. S. Parker, “Empirical comparisons of various voting methods in bagging,” in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’03*, p. 595, Washington, DC, USA, August 2003.
 - [34] S. J. MousaviRad, F. Akhlaghian Tab, and K. Mollazade, “Application of imperialist competitive algorithm for feature selection: a case study on bulk rice classification,” *International Journal of Computer Application*, vol. 40, no. 16, pp. 41–48, 2012.
 - [35] N. A. Kuchekar and V. V. Yerigeri, “Rice grain quality grading using digital image processing techniques,” *Journal of Electronics and Communication Engineering*, vol. 13, 2018.
 - [36] M. M. Agrawal and D. S. Agrawal, “Rice plant diseases detection & classification using deep learning models: a systematic review -,” *J. Crit. Rev.* vol. 7, no. 11, pp. 4376–4390, 2020.
 - [37] O. Yorulmaz, T. C. Pearson, and A. EnisÇetin, “Detection of fungal damaged popcorn using image property covariance features,” *Computers and Electronics in Agriculture*, vol. 84, 2021.
 - [38] O. Díaz, T. Ferreira, J. Rodríguez-Otero, and Á. Cobos, “Characterization of chickpea (*cicer arietinum* L.) flour films: effects of pH and plasticizer concentration,” *International Journal of Molecular Sciences*, vol. 20, no. 5, p. 1246, 2019.
 - [39] S. Sankaran, M. Wang, and G. J. Vandemark, “Image-based rapid phenotyping of chickpeas seed size,” *Engineering in Agriculture, Environment and Food*, vol. 9, no. 1, pp. 50–55, 2016.
 - [40] K. Laabassi, M. A. Belarbi, S. Mahmoudi, S. A. Mahmoudi, and K. Ferhat, “Wheat varieties identification based on a deep learning approach,” *Journal of the Saudi Society of Agricultural Sciences*, vol. 20, no. 5, pp. 281–289, 2021.
 - [41] A. Bhande and D. S. V. Rode, “Quality identification of wheat by using image processing,” *International Journal of Engineering Science and Computing*, vol. 6, 2016.
 - [42] UCI Machine Learning Repository, “UCI Machine Learning Repository: Seeds Data Set,” 2021, <https://archive.ics.uci.edu/ml/index.php>.
 - [43] G. Guo, H. Wang, D. Bell, and Y. Bi, “KNN Model-Based Approach in Classification,” in *Proceedings of the OTM Confederated International Conferences On the Move to Meaningful Internet Systems*, Amantea, Italy, October 2004.
 - [44] D. Agrawal and P. Dahiya, “Comparisons of classification algorithms on seeds dataset using machine learning algorithm,” *Compusoft*, vol. 7, pp. 2760–2765, 2018.
 - [45] H. H. Patel and P. Prajapati, “Study and analysis of decision tree based classification algorithms,” *International Journal of Computer Science and Engineering*, vol. 6, no. 10, pp. 74–78, 2018.
 - [46] B. G. Priya, “A comparison of various machine learning algorithms for wheat seed data set classification,” *An international journal of advanced computer technology*, vol. 7, no. 5, 2019.
 - [47] R. Lewis, “An Introduction to Classification and Regression Tree (CART) Analysis,” in *Proceedings of the 2000 Annual Meeting of the Society for Academic Emergency Medicine*, San Francisco, CA, USA, January 2000.
 - [48] T. N. Rincy and R. Gupta, “Ensemble learning techniques and its efficiency in machine learning: a survey,” in *Proceedings of the 2nd International Conference on Data, Engineering and Applications (IDEA)*, pp. 1–6, Bhopal, India, February. 2020.
 - [49] X. Wei and Y. Cao, “Automatic counting method for complex overlapping erythrocytes based on seed prediction in microscopic imaging,” *Journal of Innovative Optical Health Sciences*, vol. 9, no. 5, Article ID 1650016, 2016.
 - [50] D. Saravagi, S. Agrawal, and M. Saravagi, “Opportunities and challenges of machine learning models for prediction and diagnosis of spondylolisthesis: a systematic review,” *International Journal of Engineering Systems Modelling and Simulation*, vol. 12, no. 2/3, pp. 127–138, 2021.