International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning based Robust Model for Seed Germination Detection and Classification

Srinath Yasam[1], Dr. S. Anu H. Nair[2], Dr. K. P. Sanal Kumar[3*]

Submitted: 20/10/2022 **Revised:** 22/12/2022 **Accepted:** 24/01/2023

Abstract: Seed germination assessment is a quite difficult for the research team members to evaluate performance and quality. Generally, seed assessment can be performed manually, which is an error-prone, cumbersome, and time-consuming process. The typical image analysis method is not suitable for largescale germination experiments, since they frequently depend on manual adjustment of color-based threshold. Various researcher workers have projected methods to automate these processes to alleviate the manual processes in seed testing, which is extremely error-prone. Lately, image analysis technique has been used for seed detection, as they provide unbiased and quantitative measurements and can be easily automatized with minimal errors. Hence, this study designs a new Machine Learning based Robust Classification Model for Seed Germination (MLRCM-SG). The presented MLRCM-SG technique carries out the automated identification and classification of germination, to evaluate the seed quality. To attain this, the presented MLRCM-SG technique initially undergoes preprocessing in two stages namely CLAHE based contrast enhancement and median filter (MF) based noise removal. In addition, the presented MLRCM-SG technique employs Scale-Invariant Feature Transform (SIFT) technique is used in preprocessed images for the collection of feature vectors. Finally, process of classification takes place using two ML classifiers namely random forest (RF) and decision tree (DT). The experimental validation of the MLRCM-SG technique is tested on seed germination dataset and the outcome shows the remarkable performance of the MLRCM-SG algorithm compared to current approaches.

Keywords: *Germination; Seed quality; Machine learning; Artificial intelligence; classification*

1. Introduction

Germination is regarded as the primary feature of seed quality. It directly affects quality and yield. Selection of seed is critical in breeding; besides the seed quality it will be a vital part in ensuring profitable farming in addition to the seed quality [1]. The assessment of seed quality process can be generally executed probably by professional technicians, which relies upon individual knowledge and takes considerable time. Recently, scholars have modelled machine vision-oriented techniques to monitor the quality of seeds in addition to characteristics like texture, colour, and shape of seeds to automatically and quickly acquire appropriate variables and facilitate seed grading by the seed producer [2]. But either it can be an automatic classification by computers or a spontaneous judgment by technicians, the seeds grading results need to be verified by the test called typical germination test [3]. This germination test generally takes nearly two to ten days for various kinds of seeds and needs expert personnel for counting seeds manually that have germinated at some period the test should be repeated whenever the difference between the outcomes of sample does not lie in some range, which necessitates maximal efforts from skilled experts [4].

The reason behind this is the unique nature of germination tests and conventional techniques are dependent on skilled professionals for the classification of seed germinating state is a more reliable automatic system to examine the germination of seed test with more stability [5], independence, and accuracy from species type is becoming an immediate need [6]. Now, the automatic germination test monitoring system described in the research report now relies on conventional image analysis methods [7]. Conventional method of seed vigor tests is not frequently used due to complex and time-consuming process. Additionally, many of the seed tests developed by the International Seed Testing Association (ISTA) are

[1] *Research Scholar, Department of Computer Science and Engineering, Annamalai University, Chidambaram, India. Email: Srinath.yasam@gmail.com*

[2] *Assistant Professor, Department of CSE, Annamalai University, Chidambaram, India (Deputed to WPT Chennai). Email: anu_jul@yahoo.co.in*

[3] *Assistant Professor, P.G Department of Computer Science, R. V. Government Arts College, Chengalpattu, Tamil Nadu, India. Email: sanalprabha@yahoo.co.in *Corresponding Author*

evaluated manually using established procedures that differ for various crops [8]. Many of the researchers have developed the automatic method for minimize the error to reduce the number of manual steps in seed testing process. Modern image analysis techniques have been used to identify seeds because they are automated and provide quantitative, impartial measurements with little errors [9]. Some of the approaches that have been published merely image-based color-related parameters and forecast variables like area, breadth, color values, roundness perimeter and length to describe the 333seed.

Earlier, evaluation of many machines learning (ML) techniques, SVM, Naive Bayes Classifier (NBC), DTs, ANN, and KNN to compare seed germination

2. Related Works

In this section, a brief survey of the present germination classification model has been discussed. By combining automated Xray study with ML methodology, de Medeiros et al. [11] investigated a method to forecast germination ability and differentiate Jatropha curcas L. seeds by postulating the germination rate and vigor of the seedlings. Afterward the Xray test, the seeds can be yielded to physiological reviews. According to every separate seed descriptor, quality classes can be generated and LDA approaches are executed. Hyperspectral databases are introduced by Hu et al. [12] using HSI of maize seedlings with various grades of mildew, and they are then classified using ML techniques and spectral information. Primarily, the images are processed with morphological and Otus functions. All the seed's spectral characteristics can be extracted dependent on its edge, its original pixel coding, ROI, and coding. The RF techniques can be optimized utilizing the SSA that is inefficient of escape the local optimal; therefore, it can be optimized utilizing a modified reverse SSA (JYSSA) approach.

In [13], a near-infrared (NIR) HSI can be utilized for determining the possibility and vigor of naturally-aged rice seed development. In order to define the reference values of vigor and viability, typical germination testing has been conducted. Utilizing the full range spectra and feature wavelengths chosen by CNN, PCA and conventional ML techniques (SVM and LR) can be built to forecast the likelihood and potency of various variations of rice seeds under natural ageing conditions. In [14], HSIs of 3072 sugar beet seeds of similar variation can be gathered and were successively managed by binarization, extraction contour, morphology, etc. In SVM-RBF, KNN, and RF techniques can be introduced at the complete wavelength and feature

recommended higher accuracy and efficiency of ANN methods [10].

This study designs a new Machine Learning based Robust Classification Model for Seed Germination (MLRCM-SG). The presented MLRCM-SG technique performs pre-processing in two stages namely CLAHE based contrast enhancement and median filter (MF) based noise removal. In addition, the presented MLRCM-SG technique employs Scale-invariant feature transform (SIFT) technique used in pre-processed images for the collection of feature vectors. Finally, two ML classifiers such as RF and DT are used in the classification of seed germination process. The experimental validation of the MLRCM-SG technique is tested on seed germination dataset.

wavelength correspondingly for predicting the sugar beet seeds germination. In [15], we can obtain the NIR-HIS of 3072 beetroot seeds and the corresponding germination status. Based on the study of node information gain of the developed tree approach and the physical properties of beetroot seeds, 15 features may recovered to reduce the dimensions of data assessment and to be used with RF, Light GBM and SVM (RBF) classifier approaches.

In [16], a new method for seed quality classifier was projected. The authors established classification approaches utilizing Fourier transform NIR (FT-NIR) spectroscopy and Xray imaging methods for predicting vigor and seed germination. Radiographic image and FT-NIR spectroscopy dataset can be attained in separate seeds, and these methods can be generated depending upon the subsequent techniques: NB, LDA, PLS-DA, and SVM with radial basis (SVM-r) kernel RF are all examples of discriminant analysis. Using Xray imaging techniques to examine the inner characteristic (air space and endosperm) of three varieties of watermelon seeds, conventional DL and ML algorithms may be approximated in [17]. The Fisher objective function and sequential forward selection (SFS) approach can be utilized as a searching approach and feature optimized.

3. The Research Methodology

In this research methodology, an automated germination classification model, named MLRCM-SG technique, which helps to determine the seed quality. The presented MLRCM-SG technique follows a series of sub-processes namely CLAHE based contrast enhancement, SIFT feature extraction, MF based noise elimination, and ML based classification. Fig. 1 demonstrates the workflow of MLRCM-SG approach.

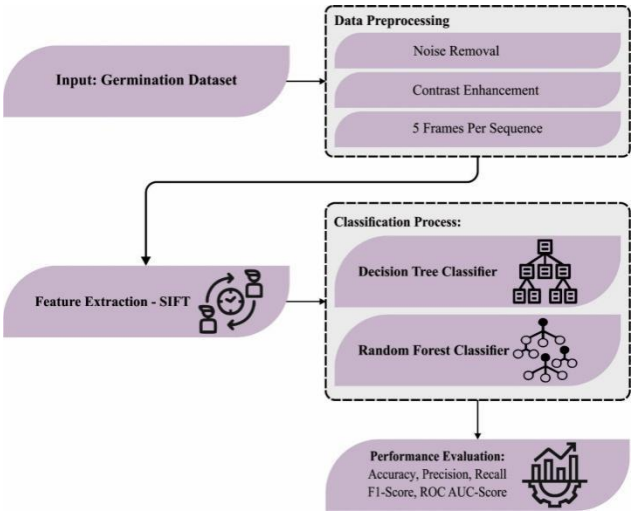

Fig. 1. Scheduling of MLRCM-SG approach

3.1. Image Pre-processing

At the first stage, two stages of image pre-processing were carried out: noise removal using MF and contrast enhancement using CLAHE. An MF with the kernel size of $n \in n$ is used for image enhancement [18]. Image pixel was filtered with $n \in n$ kernel around every pixel, which replace the central pixel values with the central values of each pixel existing in the kernel, and the mathematical expression can be given as follows:

$$H_{med}(X,Y) = \left[k_{n,n}(i,j) \otimes H_{enh}[X+i, Y+j]\right] \quad (1)$$

The kernel size was set to 3 to prevent over-filtering. Destructive interference suppression can be formulated by the subsequent expression.

$$H_c(X,Y) = \left(H_{enh}(X,Y), H_{med}(X,Y)\right) \quad (2)$$

Let A_c be the constructive interference image. Constructive interference suppression done by using the MF is employed to H_c for removing the remaining extreme single pixel. CLAHE algorithm can be expressed based on splitting the images into non-overlapping regions of approximately equivalent size. In order to splitting the image evenly by 8 in all directions, the number of regions is often chosen equally to 64 which allowing for better statistical analysis. This partition leads to three dissimilar groups of regions. The class of corner regions (CR) comprises four regions, is. The class of boarder region (BR) comprises 24 regions. The class of inner region (IR) comprises all the residual 36 regions. Every region on the image boarder, excluding the corner region, belongs to these classes. In the presented method, first, histogram of all the regions is evaluated. Next, based on the preferred limit for contrast expansion, a clip limit for clipping histogram can be attained. Then, every histogram is redistributed so that its height doesn't exceed the clip limit. Lastly, cumulative distribution function, CDF, of the consequential contrast limited histogram is defined for grayscale mapping. In this method, pixels are mapped by linearly integrating the result from the mapping of the four adjacent regions.

3.2. Extract Features using SIFT

This research work, a set of feature vectors is created using the SIFT approach. SIFT is a region descriptor or detector extracting image features that are partially unvarying to the changes in the three-dimensional camera view (affine transformation) and illumination as well as invariant to image rotation and scaling [19]. Feature detection can be done by a staged filtering method, based on the experimental and theoretical outcomes. The hypothesis is that only potential SIFT image space can be provided by the Gaussian function. The scale-space can be determined by:

$$L\ a,b,)\ G\ a,b,\ \ I\ a,b \quad (3)$$

Lowe introduced a technique for the recognition of SIFT keypoint. SIFT extracts keypoint as the maximal response of Difference-of-Gaussian (DoG) operation that is calculated from the difference of two neighboring scales s divided by the k constant multiplicative factor

$$\begin{aligned} &D\ (\ a\ ,b\ ,\sigma\ ) \\ &=\ (\ G\ (\ a\ ,b\ ,k\sigma) - G\ (\ a\ ,b\ ,\sigma\ )\ ) \\ &*\ I\ (\ a\ ,b\ ) \\ &= L\ (\ a\ ,b\ ,k\sigma\ ) - L\ (\ a\ ,b\ ,\sigma\ ) \end{aligned} \quad (4)$$

In contrast to the Lindeberg algorithm, the major benefit of Lowe's detector is the computation cost that is minimized. This is because the scale-normalized Laplacian of Gauss is replaced with the difference of Gaussian functions, viz., calculated by the image subtraction. Feature point is identified in the DoG as the local minima and maxima of $D(a,b,\sigma)$. In a predetermined image scale, all the image points are compared with the 9 neighbors and the 8 neighboring pixels in the above and below scale. Local minima or maxima are categorized as keypoint. For selecting a stable keypoint, the local extrema value of $D(a,b,\sigma)$ needed to be maximum when compared to the threshold (Th_key). Brown proposed a local detection technique that can be refined by sub-pixel localization. Local maxima are recognized by fitting the $3D$ quadratic function $D(a,b,\sigma)$ to the local sample point using Taylor expansion. The robustness of local detection can be enriched by eliminating keypoint with lower contrast. The $D(a,b,\sigma)$ scale-space calculated at the sample point is, especially, beneficial to remove low-contrast extrema.

The DoG scale-space provides keypoint that is invariant to scale transform, however, it is sensitive to rotation. To

prevent this problem during feature matching, the SIFT detector allocates a canonical orientation at all the keypoints, by using the local radiometric property of the neighboring pixel. The image matching between keypoint extracted in the stereo pair can be performed by means of SIFT descriptor. The matching aims at "describing" a local image region around keypoint. This technique uses histogram to characterize various features of the shape and appearance. The considerable benefit of these operators is their stability against smaller shifts in the local geometry, which arises from homographic or affine projection.

3.3. Germination Classification

For germination classification process, two ML classifiers namely RF and DT are used in this study.

3.3.1. RF Model

Each tree in the RF classifier gives one vote to the better-known class for classifying the input vector. They are generated using random vector samples taken one at a time from the input vector [20]. The RF classifier uses feature combinations or randomly selected features at every node to produce the tree. By randomly permuting N samples, a training dataset is created through the process of bagging. For every feature or feature combination chosen, N represents the size of the unique training dataset. The number of examples (pixels) is categorized by taking the more commonly known class from each tree predictor in the forest. A pruning mechanism and the FS measure are needed for the DT model. There are numerous methods to the FS applied for DT induction and most techniques directly allocate a quality measure to the attributes. The more commonly known feature selection measure in DT induction is Gini Index and Information Gain Ratio criterion. The convenience use of Gini coefficient as a FS measure which takes advantages of RF classifier to measure the class related impurities. The Gini index is created for the given training set T by randomly selecting each case (pixel) and taking into account that they are members of a particular class C_i, the Gini coefficient is formulated by:

$$\sum \sum_{j \neq i} (f(C_i, T)/|T|)(f(C_i, T)/|T|) \quad (5)$$

From the expression, f (C_i, T) / | T | indicate the possibility that the particular example falls under the category C_i.

When the tree reaches full depth on the novel training data set using feature fusion. This mature tree has not been trimmed. This is the RF classifier's biggest disadvantage in comparison to other DT models. As the overall number of trees increases, the observation error frequently encounters without overloading and trimming the tree. The quantity of trees to be grown and the number of characteristics used at each node to create a tree are two user-determined criteria required for creating RF classifiers. At all the nodes, only selected feature was searched for the best split. Therefore, the RF classification comprises N trees. To categorize a new dataset, all the cases of the data set are passed down to every N tree. The forest selects a class that has maximum of N votes.

3.3.2. DT Model

A new DT classifier named C4.5 is an appropriate selection for medical data classification [21]. These hierarchical classifiers play a crucial role in the efficient classification of datasets. Quinlan in 1993 proposed a C4.5 algorithm that is extension of ID3 (Iterative Dichotomiser 3) algorithm and defined it as a typical ML algorithm for supervised classification and extensively applied it as a data analysis tool for dissimilar regions. The tree construction initiates at the root node and is extended in a top-down algorithm. In the DT model, every attribute is measured by means of a statistical estimation for deciding how precisely it could classify the training instance. The key attribute is selected as the root node of the tree. Next, the training dataset is stored based on the successor node. This searching technique is reiterated for selecting the key attribute for the training instances. This procedure forms a greedy search for the DT model and it never backs track to analyze the preceding node.

Once every sample is assigned to the one node that belongs to a similar class at that time then tree stops to add the node to it. A new node is originated from the parent tree only if there are sufficient samples left from the sorting. Tree pruning is widely implemented to prevent overfitting of data afterward the full construction of tree. The C4.5 approach makes use of a statistical test to allocate the attribute to all the nodes according to the entropy measure. Based on the maximum information gain ratio the attribute was chosen to construct the tree as follows:

$$Gain\ Ratio(A,S) = \frac{Gain(A,S)}{Spilt\ Information(A,S)}, \quad (6)$$

Where $Gain(A,S) = Ent(S) - \sum_{a \in A} \frac{|S_a|}{S} Ent(S_a)$ and

$$Spilt\ Information\ (A,S) = -\sum_{a \in A} \frac{|S_a|}{S} \frac{|S_a|}{S}.$$

The information gain ratio is directly computed for discrete valued attributes; at the same time, continuous valued

attributes must be discretized for the computation of information gain ratio.

4. Experimental Validation

The MLRCM-SG model's germination categorization results are evaluated in this section using a dataset [1] that consists of photos divided into three classes. The MLRCM-SG model has taken five images per sequence. In Table 1. the details related to the dataset are illustrated. In Fig. 2, the sample images are shown.

Table 1 Dataset specifications

Descriptions	Pennisetum Glaucum	Secale Cereale	Zea Mays
Total_Images	7954	7695	8148
Total_Sequences	82	81	84
N_Images_Per_Seq	6	5	6
Total_Selected_Images	492	405	504

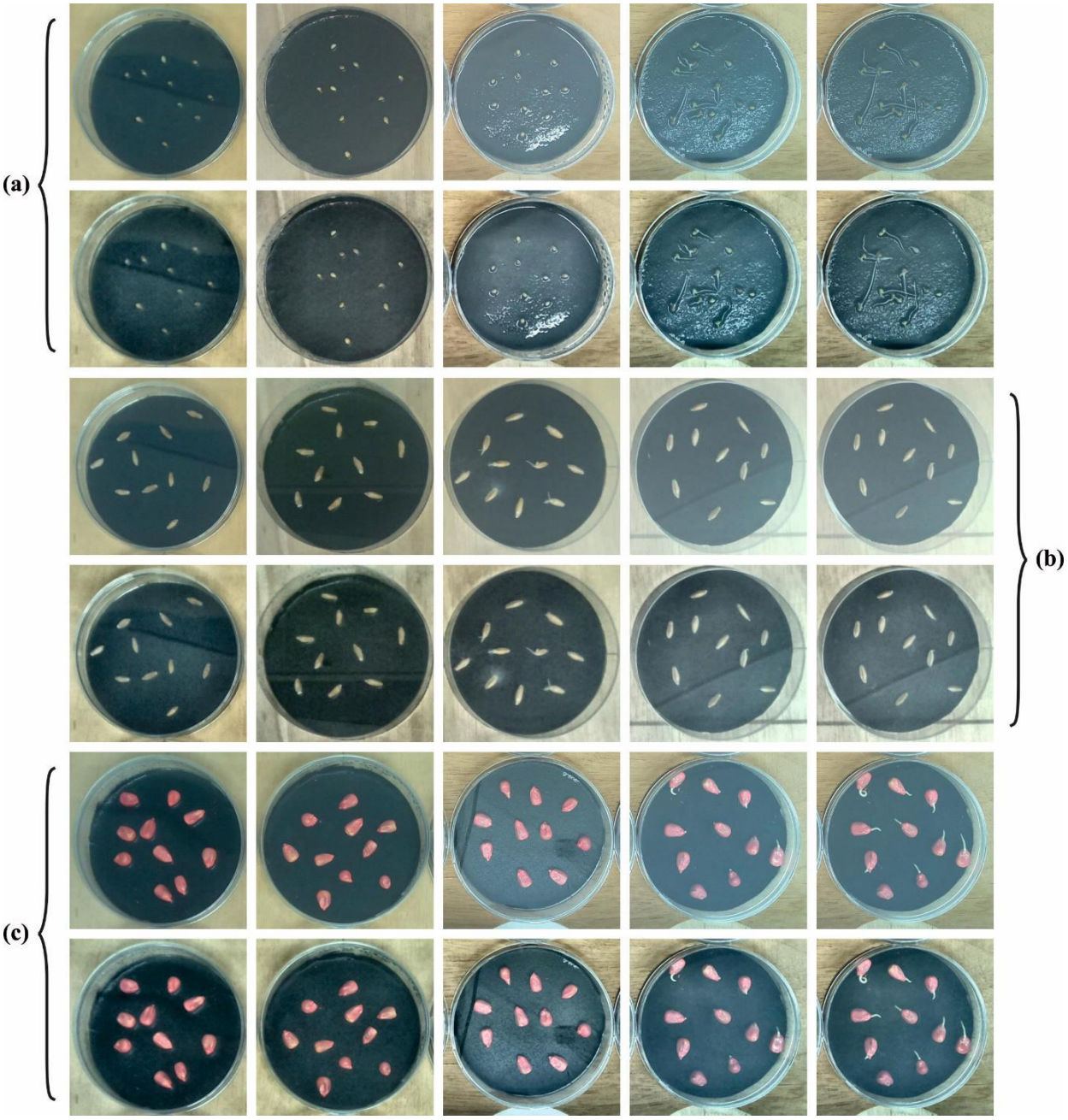

Fig. 2. Model Images a) represents Pennisetum Glaucum b) represents Secale Cereale c) represents Zea Mays

The presented model's attributes are shown in Fig. 3 on five photos each sequence. The traits taken from samples of Pennisetum Glaucum are displayed in Fig. 3a. Similar traits from samples of Secale Cereale are shown in Fig. 3b. Additionally, Fig. 3c presents the features derived from Zea Mays data sample.

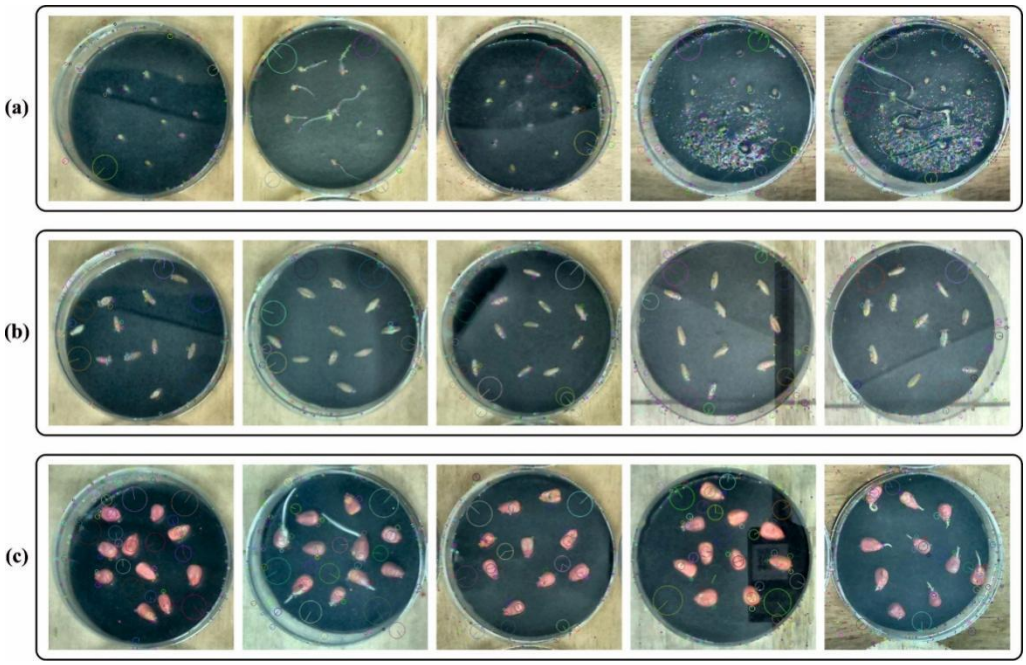

Fig. 3. Feature Extraction Visualization phase. Phase (a) Pennisetum Glaucum Phase (b) Secale Cereale Phase (c) Zea Mays

The DT approach using TS and TR data are shown in Figure 4. The Confusion Matrix (CM) provided by the DT technique using TR data is shown in Fig. 4a. According to the graph, 301 instances were detected in PG, 230 in SC, and 340 in ZM using the DT method. Next, the CM provided by the DT method under TS data is shown in Fig. 4b indicates 129 PG instances, 112 SC and 139 ZM cases can be detected and outperformed by the DT system. A PR study analysis of the DT algorithm under TS and TR data is also defined in figs 4c-4d. Numerically, the DT strategy achieved the highest PR scores in three different classes. Finally, utilising TR and TS data, Figs. 4e and 4f demonstrate the ROC analysis of the DT approach. The graph demonstrated that three classes with higher ROC values saw proficient results using the DT approach.

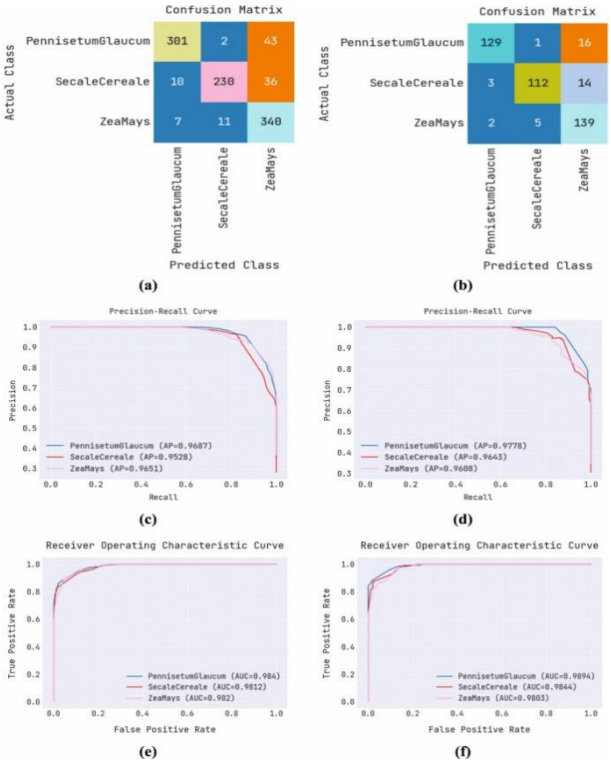

Fig. 4. Training Phase of a DT Classifier: a) CM c) Curve for PR e) ROC Testing Phase b) CM d) Curve for PR f) ROC

Table 2. shows how the DT approach classifies sprouting using TS and TR data.

Decision Tree Algorithm		
Metrics	**Trained Data**	**Test Data**
Accuracy	88.88	90.26
Precision	90.15	91.14
Recall	88.43	90.13
F1-Score	88.94	90.36
ROC Score	98.24	98.47

Table 2 and Fig 5 both display the total seeding classification results for the DT model beneath the TS and TR data. The outcomes demonstrated that the DT algorithm produced beneficial outcomes for both TS and TR data. For example, the DT model uses TR data to provide results. $accu_y$ of 88.88%, $prec_n$ of 90.15%, $reca_l$ of 88.43%, $F1_{score}$ of 88.94%, and ROC_{score} of 98.24%. Meanwhile, with TS data, the DT approach has offered $accu_y$ of 90.26%, $prec_n$ of 91.14%, $reca_l$ of 90.13%, $F1_{score}$ of 90.36%, and ROC_{score} of 98.47%.

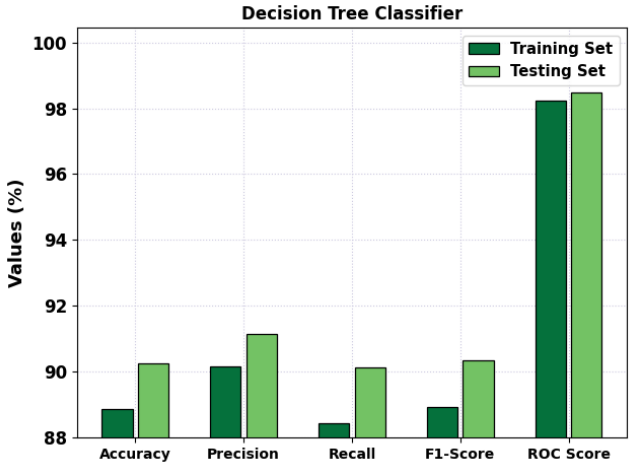

Fig. 5. Results of the DT method for classifying germination usingg TR and TS values

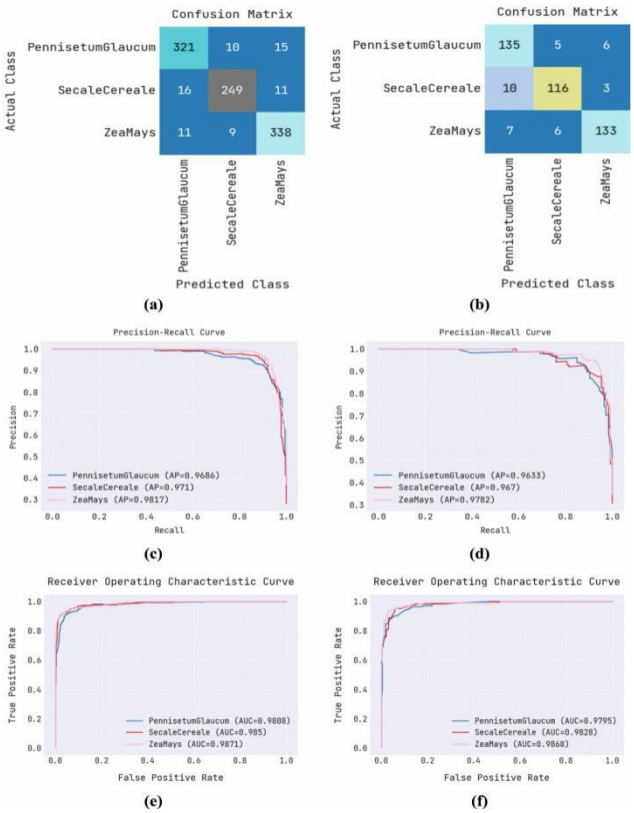

Fig. 6. RF algorithm a) CM c) Curve for PR e) ROC Testing Phase b) CM d) Curve for PR f) ROC

The classification result of the RF approach using TS and TR data is shown in Fig. 6. The CM provided by the RF technique using TR data is shown in Fig. 6a. According to the figure, 249 times under SC, 338 times under ZM and 321 times under PG have been discovered using the RF technique. Additionally, Fig. 6b represents the CM made available by the RF algorithm for TS data. According to the graph, the RF method has found 133 ZM instances, 116 SC instances, and 135 PG instances. The PR investigation of the RF method under TS and TR data is also defined in Figs. 6c and 6d. According to the statistics, the RF method has improved PR performance across three class labels. Finally, ROC analysis of the RF approach employing TS and TR statistics is shown in Figs. 6e and 6f. According to the fig, by employing the RF method, produced effective results in three classes with higher ROC values.

Table 3 Results of the RF approach for classifying germination using TR and TS

Random Forest Algorithm		
Metrics	**Trained Data**	**Test Data**
Accuracy	92.65	91.21
Precision	92.67	91.27
Recall	92.47	91.16
F1-Score	92.56	91.20
ROC Score	98.43	98.30

Table 3 and Fig.7 represents the average germination categorization results of the RF technique are provided under TS and TR statistics. The results showed that RF technique was effective on both TR and TS data. For sample, with TR data, the RF system has offered $accu_y$ of 92.65%, $prec_n$ of 92.67%, $reca_l$ of 92.47%, $F1_{score}$ of 92.56%, and ROC_{score} of 98.43%. In the meantime, with TS data, the RF approach has offered $accu_y$ of 91.21%, $prec_n$ of 91.27%, $reca_l$ of 91.16%, $F1_{score}$ of 91.20%, and ROC_{score} of 98.30%.

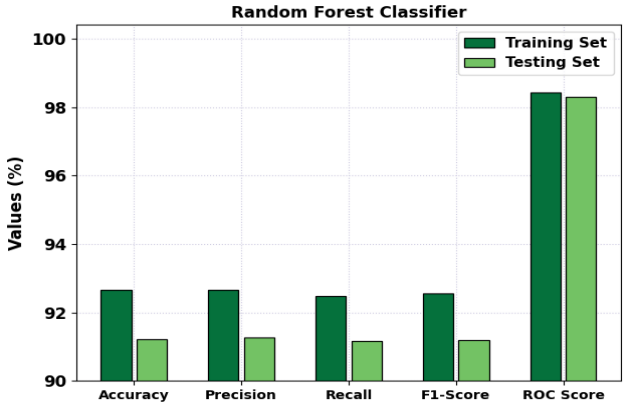

Fig. 7. Germination classification outcome of RF method under TS and TR data

Finally, a comparative germination classification performance of the MLRCM-SG model with other classifiers is provided in Table 4 and Fig. 8 [5]. According to the experimental results, the GNB model has a lower $accu_y$ of 69.93%. Following that, the Linear SVM, Fine KNN models, K-NB, Weighted KNN and Quadratic-SVM models provided slightly improved results of 74.71%, 79.55%, 74.90%, 73.64% and 77.18%, respectively.

Table 4 Comparison of the MLRCM-SG approach's accuracy with certain other classifiers

Techniques	Accuracy (%)
MLRCM-SG	92.65
GNB Model	69.93
Linear SVM	74.71
Fine KNN	79.55
Kernal-Naïve Bayes	74.90
Weighted KNN	73.64
Quadratic SVM	77.18
Linear Discriminant Model	86.88
Subspace Discriminant Ensemble	91.60

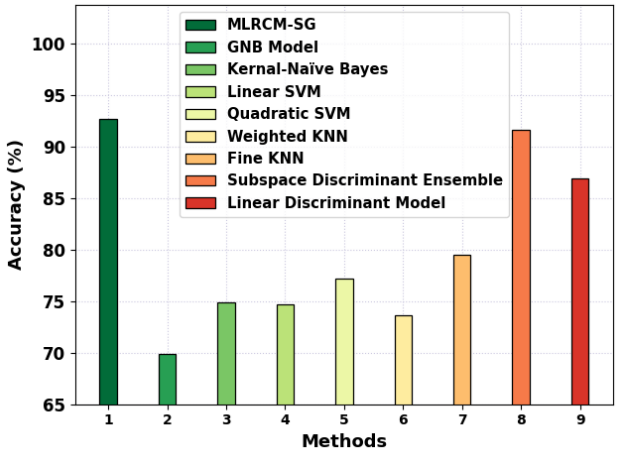

Fig. 8. $Accu_y$ analysis of MLRCM-SG approach with other classifiers

Meanwhile, the linear discriminant model has reached reasonable $accu_y$ of 86.88%. Though the subspace discriminant ensemble model has obtained considerable $accu_y$ of 91.60%, the MLRCM-SG model has gained better outcomes with $accu_y$ of 92.65%. These results assured the improved germination classification performance of the MLRCM-SG model.

5. Conclusion

In this study, we have designed an automated germination classification model, named MLRCM-SG technique, which helps to determine the seed quality. Primarily, the presented MLRCM-SG technique performed preprocessing in two stages namely CLAHE based contrast enhancement and MF based noise removal. The feature vectors are collected from the pre-processed images using MLRCM-SG technique and the SIFT technique. Finally, two ML classifiers, RF and DT, are used in the classification of seed germination procedure. The experimental validation of the MLRCM-SG technique is tested on seed germination dataset and the outcome shows the remarkable performance of the MLRCM-SG algorithm compared to current techniques. Therefore, the MLRCM-SG technique can be employed for accurate germination classification. In the future, DL techniques can be employed to enhance the classification performance of the MLRCM-SG technique.

References

[1] Genze, N., Bharti, R., Grieb, M., Schultheiss, S.J. and Grimm, D.G., 2020. Accurate machine learning-based germination detection, prediction and quality assessment of three grain crops. *Plant methods*, *16*(1), pp.1-11.

[2] Colmer, J., O'Neill, C.M., Wells, R., Bostrom, A., Reynolds, D., Websdale, D., Shiralagi, G., Lu, W., Lou,

Q., Le Cornu, T. and Ball, J., 2020. SeedGerm: a cost-effective phenotyping platform for automated seed imaging and machine-learning based phenotypic analysis of crop seed germination. *New Phytologist*, *228*(2), pp.778-793.

[3] Nehoshtan, Y., Carmon, E., Yaniv, O., Ayal, S. and Rotem, O., 2021. Robust seed germination prediction using deep learning and RGB image data. *Scientific reports*, *11*(1), pp.1-10.

[4] Yang, U., Oh, S., Wi, S.G., Lee, B.R., Lee, S.H. and Kim, M.S., 2021. Classification of Germination Images of Pear Pollen Using Random Forest and Convolution Neural Network Models. *IEEE Access*, *9*, pp.45993-45999.

[5] Durai, S. and Mahesh, C., 2021. Research on varietal classification and germination evaluation system for rice seed using hand-held devices. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, *71*(9), pp.939-955.

[6] Yasam, S., Nair, S.A.H. and Kumar, K.P., 2022. Supervised learning-based seed germination ability prediction for precision farming. *Soft Computing*, pp.1-12.

[7] Aasim, M., Katırcı, R., Akgur, O., Yildirim, B., Mustafa, Z., Nadeem, M.A., Baloch, F.S., Karakoy, T. and Yılmaz, G., 2022. Machine learning (ML) algorithms and artificial neural network for optimizing in vitro germination and growth indices of industrial hemp (Cannabis sativa L.). *Industrial Crops and Products*, *181*, p.114801.

[8] Valente, J., Kooistra, L. and Mücher, S., 2019. Fast classification of large germinated fields via high-resolution UAV imagery. *IEEE Robotics and Automation Letters*, *4*(4), pp.3216-3223.

[9] Durai, S., 2021. Labelled Image Dataset Preparation for Rice Seed Germination Prediction and Variety Classifictaion using Low Cost Devices. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(9), pp.245-249.

[10] Awty-Carroll, D., Clifton-Brown, J. and Robson, P., 2018. Using k-NN to analyse images of diverse germination phenotypes and detect single seed germination in Miscanthus sinensis. *Plant methods*, *14*(1), pp.1-7.

[11] de Medeiros, A.D., Pinheiro, D.T., Xavier, W.A., da Silva, L.J. and dos Santos Dias, D.C.F., 2020. Quality classification of Jatropha curcas seeds using radiographic images and machine learning. *Industrial Crops and Products*, *146*, p.112162.

[12] Hu, Y., Wang, Z., Li, X., Li, L., Wang, X. and Wei, Y., 2022. Nondestructive Classification of Maize Moldy Seeds by Hyperspectral Imaging and Optimal Machine Learning Algorithms. *Sensors*, *22*(16), p.6064.

[13] Jin, B., Qi, H., Jia, L., Tang, Q., Gao, L., Li, Z. and Zhao, G., 2022. Determination of viability and vigor of naturally-aged rice seeds using hyperspectral imaging with machine learning. *Infrared Physics & Technology*, *122*, p.104097.

[14] Yang, J., Sun, L., Xing, W., Feng, G., Bai, H. and Wang, J., 2021. Hyperspectral prediction of sugarbeet seed germination based on gauss kernel SVM. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *253*, p.119585.

[15] Zhou, S., Sun, L., Xing, W., Feng, G., Ji, Y., Yang, J. and Liu, S., 2020. Hyperspectral imaging of beet seed germination prediction. *Infrared Physics & Technology*, *108*, p.103363.

[16] Medeiros, A.D.D., Silva, L.J.D., Ribeiro, J.P.O., Ferreira, K.C., Rosas, J.T.F., Santos, A.A. and Silva, C.B.D., 2020. Machine learning for seed quality classification: An advanced approach using merger data from FT-NIR spectroscopy and X-ray imaging. *Sensors*, *20*(15), p.4319.

[17] Ahmed, M.R., Yasmin, J., Park, E., Kim, G., Kim, M.S., Wakholi, C., Mo, C. and Cho, B.K., 2020. Classification of watermelon seeds using morphological patterns of X-ray imaging: a comparison of conventional machine learning and deep learning. *Sensors*, *20*(23), p.6753.

[18] Bosakova-Ardenska, A., Kutryanska, M., Boyanova, P. and Panayotov, P., 2022, August. Application of images segmentation and median filter for white brined cheese structure evaluation. In *AIP Conference Proceedings* (Vol. 2570, No. 1, p. 020014). AIP Publishing LLC.

[19] Lokku, G., Reddy, G.H. and Prasad, M.G., 2022. Optimized scale-invariant feature transform with local tri-directional patterns for facial expression recognition with deep learning model. *The Computer Journal*, *65*(9), pp.2506-2527.

[20] Singh, B., Sihag, P. and Singh, K., 2017. Modelling of impact of water quality on infiltration rate of soil by random forest regression. *Modeling Earth Systems and Environment*, *3*(3), pp.999-1004.

[21] Abpeykar, S. and Ghatee, M., 2019. An ensemble of RBF neural networks in decision tree structure with knowledge transferring to accelerate multi-classification. *Neural Computing and Applications*, *31*(11), pp.7131-7151.