Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 26, Issues 2, pp. 940–949, DOI 10.12694/scpe.v26i2.3924

CLASSIFICATION OF ROYAL DELICIOUS APPLES USING HYBRID FEATURE SELECTION AND FEATURE WEIGHTING METHOD BASED ON SVM CLASSIFIER

SANDEEP KAUR[*], MANOJ KUMAR SACHAN[†] AND ASHWANI KUMAR AGGARWAL[‡]

Abstract. Fruit safety is a critical component of the global economy, particularly within the agricultural sector. There has been a recent surge in the incidence of diseases affecting fruits, leading to economic setbacks in agriculture worldwide. Conventional manual assessment methods are laborious, prompting the exploration of automated computerized techniques for evaluating fruit quality. This research presents a novel method for assessing the quality of golden delicious apples. A dataset comprising 1256 apple images was gathered under controlled conditions. Afterward Feature extraction focuses on texture features like LBP, GLCM, GLDM, DTF, and Gabor features, color features, and shape and size features. A total of 18654 features are extracted and normalized using z-score. A hybrid method for feature selection and weighting involves the mRMR algorithm to eliminate redundant features and the Sine Cosine Optimization Algorithm for feature weighting, enhancing classification performance. The SVM machine learning technique, augmented with optimized features, yielded a 10.53% improvement in accuracy compared to SVM alone. Validation against state-of-the-art methods using Friedman's mean rank test underscored the statistical significance of this approach across various metrics.

Key words: Fruit grading, normalization, feature selection, machine learning

1. Introduction. In the last decades, machine learning plays a vital role in the horticulture field. Various tasks like defect detection in products [33] , reliable yield prediction [19] , growth monitoring [36] , quality detection [27] , soil testing [8, 35] , etc. are performed with the help of machine learning techniques. The machine learning advances are mainly directed toward the fruit quality assessment [24, 3, 16]. To attain this goal two methods are used. In the first one,the grading is performed manually which is a too tedious, labor-intensive, inconsistent, and costly method [12] . The second method is grading performed with the help of machines. These machines will grade the apples based on size only, which is not an efficient method because other features such as color, texture, etc. also play a significant impact on the grading process.

To overcome these drawbacks machine learning is growing in the apple quality estimation field.The production of apples is the key player among all the fruits. India ranked fifth as a producer of apples all over the world [11]. There is a techniques in which a robotic system for fruit picking and grading, using computer vision for detection and a 4-DOF robotic arm for handling is developed. It enhances efficiency, reduces labor, and processes fruit in an average of 15 seconds for poor quality and 21 seconds for good quality, using Python and OpenCV [6]. The use of effective machine learning algorithms to combine image-based categorization approaches is at the leading edge of apple grading. Because of the uneven distribution of light on the apple's surface and the similar appearance of defects with stems and calyx, automatically grading the apples remains a difficult operation [25]. There is a study in which investigation is done that how color-balancing methods improve the classification of physiological disorders in apples under various lighting conditions using pre-trained CNN models. The ResNet50V2 model achieved 0.949 accuracy with green light sharpness data. Enhanced image quality was confirmed by improved PSNR, MSE, and SSIM metrics, demonstrating that color balancing significantly boosts classification performance.

Efficient and accurate classification is crucial for various applications, including quality assessment, sorting,

[*]Department of Computer Science and Engineering, Sant Longowal Institute of Engineering and Technology, Longowal, 148106, Punjab, India

[†]Department of Computer Science and Engineering, Sant Longowal Institute of Engineering and Technology, Longowal, 148106, Punjab, India

[‡]Department of Electrical and Instrumentation Engineering, Sant Longowal Institute of Engineering and Technology, Longowal, 148106, Punjab, India

and marketing strategies Several scientists introduced various methods to grade the apple based on physical qualities such as defect, color, texture, shape, and size are considered important grading parameters [48, 45]. The classification of apples using feature selection and feature weighting techniques has been a subject of considerable research interest in agricultural and food industries. Feature selection is pivotal in enhancing classification accuracy by identifying the most relevant and discriminative features from input data. In apple classification, diverse methods such as filter, wrapper, and embedded approaches have been explored to accomplish this task. These methods focus on reducing the feature space dimensionality while preserving essential discriminatory information, thus boosting classification algorithm performance embedded [41, 43] . Conversely, feature weighting techniques allocate suitable weights to individual features based on their significance in classification tasks. This approach mitigates the influence of imbalanced or noisy features, thereby enhancing classifier performance. Notable methods include ReliefF [22] , and Chi-square [18] , which prioritize features according to their relevance and contribution to classification accuracy.

Lee et al. introduced color based technique for the determination of the apple maturity stage as well as the defective surface using Open CV software [24] .Yuan et al. also developed color based technique. For classification two methods were used one is SVM and the second one is the particle swarm optimization (PSO) algorithm [49] . Suresha M. et al. developed a color- based technique to classify the apples. The classification was performed by using an SVM classifier. A total of 90 images were used for experimentation [39] . Guo et al. introduced a back-propagation neural network method to classify multiple fruits and vegetables. The main focus of the classification is on color and texture. [15] . Bhatt and pant introduced Artificial Neural Network(ANN) based technique to classify the apples using size, shape, and color features. A total of 199 images were used to train the ANN [5] . Sofu et al. introduced A decision tree classification method based on color, weight,and shape features that were utilized for Granny Smith, Golden, and Starking Delicious apples. A total of 732 images are collected in which defect detection in granny smith and golden delicious was easy in comparison to starking delicious [38]. Moallem et al. introduced a technique for golden delicious apple grading based on surface features such as defect, texture, geometrical, and color features. The 120 images were collected for experimentation [30] . Ali and Thai introduced a technique for apple and mango grading by considering defects as the main features [2]. Bhargava and Bansal came up with fruit quality estimation techniques with the help of four classifiers k-nearest neighbor (k-NN), the sparse representative classifier (SRC), support vector machine (SVM), and artificial neural network (ANN)[4].

In recent literature, researchers have introduced machine learning methods for grading apples by combining multiple types of features. While this fusion can enhance the comprehensiveness of the analysis, it often complicates the grading process. The presence of redundant or irrelevant features can adversely affect the performance of classifiers. Moreover, the existence of outliers in the feature set can further diminish overall performance. These challenges serve as the primary motivation for our proposed approach, which aims to tackle these issues to boost classifier efficiency. Our method involves refining the feature set to eliminate redundancy and irrelevance while effectively managing outliers. By addressing these aspects, we aim to streamline the grading process, improve accuracy, and offer a more robust and reliable solution for apple grading using advanced machine learning techniques.

2. Proposed Methodology. In this work, a methodology is proposed to improve the classification of apples. The proposed approach is divided into six phases. In the first phase, the apple images are captured using the image acquisition mechanism. Then, the regions of the images are segmented using the image segmentation method in the second phase. In the next phase, the different features are extracted from the images based on the texture, color, shape and amp; size. A data normalization method is used to prevent larger numeric feature values from dominating smaller numeric feature values. The major objective of the normalization is to reduce the bias of those characteristics in pattern classes that contribute numerically more than others. After normalizing the features, the redundant and irrelevant features are removed with the help of a feature selection method. The reduced feature set is used for the classification of the apples in the last phase. The block diagram of the proposed approach has been illustrated in Figure 2.1.

2.1. Image acquisition system. The image acquisition system has been composed of 2 charged coupled devices (CCD) color cameras (SONY cyber shot DSC-W800). The system provides 2492x2492 color images. The lighting system has been composed of two T-shaped fluorescent tubes at the top of the square- shaped

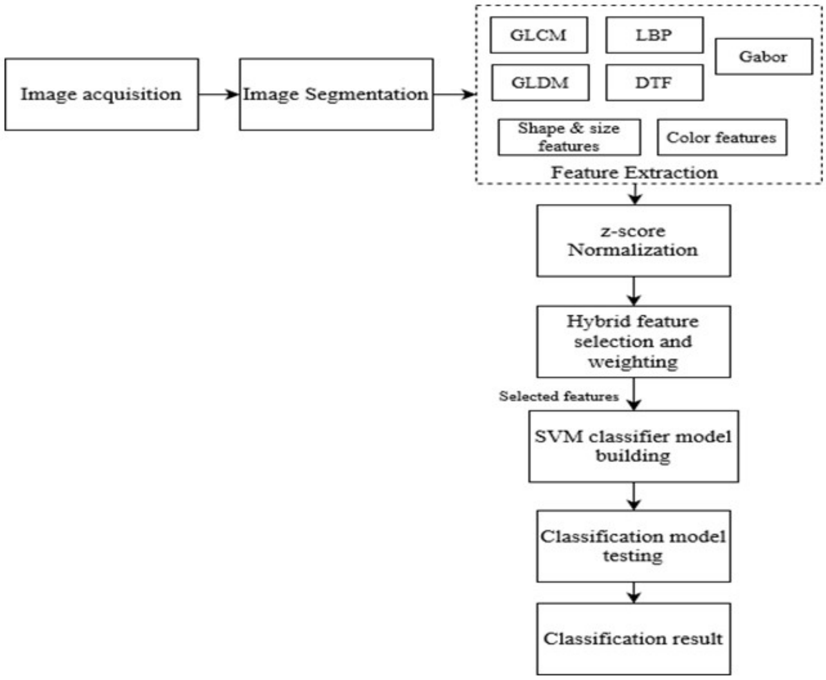

Fig. 2.1: Block diagram of the proposed approach

chamber. The chamber contains two holes for cameras, one at the top of the chamber and one at the left wall of the chamber. Reflecting surface protection has been placed between the fluorescent tube and apple to avoid direct light. Six images have been collected by rotating the apple thirty-degree horizontally using the top camera and two images from the side camera. The dataset has been created from 158 royal delicious apples collected from a Kashmir farm. A total of 1264 images have been captured which contain 256 images of grade A, 336 images of grade B, grade C, and grade D. The sample of images captured are shown in figure 2.2.

2.2. Image Segmentation. In the proposed approach, segmentation is used as a pre-processing method. The segmentation methods are used to divide the images into different isolated regions. In this approach, to divide the image into two regions (fruit area and background), the segmentation of 1256 images is performed by using the advanced photo editor Adobe Photoshop CC which is an image processing tool. The segmentation image result samples are shown in Figure 2.3.

2.3. Feature Extraction. In the proposed approach, different features are extracted from the apple images for the classification of apples into different classes. The features are extracted to learn the pattern between the different samples of apples related to the different classes. In this approach, features are extracted based on the texture, color and shape and amp; size of the apples. The details of the features are given below:

Texture Features. Image texture is a collection of two-dimensional arrays that have been generated to measure the visual texture of an image [32]. It gives information about how colors or intensities are arranged spatially in a region of an image. In this approach different texture features are extracted from collected images of apples. The texture which is used in this approach is linear binary patterns (LBP) [31] , Gray-Level Co-occurrence Matrices (GLCM) [17] , Gray- Level Difference Method (GLDM) [47] , difference theoretic features (DTF) [40] , Gabor [20] features are extracted for the classification. The description of texture feature extraction method is as follows:

Fig. 2.2: Different grades of apple: (a) Grade A, (b) Grade B, (c) Grade C, (d) Grade D

Fig. 2.3: Different grades of apple after segmentation: (a) Grade A, (b) Grade B, (c) Grade C, (d) Grade

Linear Binary Patterns (LBP). The texture features of the apple images are calculated using the LBP feature extraction method.The squared neighbourhood pixel values are computed in traditional LBP methods. Scale, illumination, and rotation invariance are computed in order to get all of the texture characteristics, then the magnitude vector in the difference vector is also taken into consideration [44] . The features can be calculated as follows:

Let pixel c_α be in the center of the neighbourhood image pixels.The P_x is considered as the index for the image pixel as $(g_0, \ldots, g_{P_{x-1}})$. Each binary pixel value is multiplied by a binomial factor to propduce the pixel's

LBP feature.

$$LBP_{PX} = \sum_{P_x=0}^{P_x=1} d(g_{p_x} - g_a), \tag{2.1}$$

where

$$d(n) = \begin{cases} 1, n \geq 0, \\ 0, n < 0. \end{cases} \tag{2.2}$$

Gray-Level Co-occurrence Matrices (GLCM). The GLCM features were based on energy, homogeneity, contrast, correlation, mean, entropy, and maximum probability etc. The GLCM features are calucated with the thirty-four statistical features for the different orientations $(0^0, 45^0, 90^0, 135^0)$.

Gray-Level Difference Method (GLDM). This feature extraction method is based on two pixels that were separated by a particular displacement δ and had an absolute difference in the grey levels. The motion vector and the probability density functionis defined as follows:

$$\delta = (\Delta x, \Delta y),\ S_\delta(x, y) = |S(x, y) - S(x + \Delta x, y + \Delta y)| \tag{2.3}$$

$$D(i|\delta) = Prob[S_\delta(x, y) = i] \tag{2.4}$$

where $\Delta x, \Delta y$ are the parameters of the methods, the input image is represented by $S_\delta(x, y), x, y,$ are positions in image $S_{\delta(x,y)}$ with $1 \leq x \leq M$ and $1 \leq y \leq N$. Then, a feature vector is generated by concatenating the mean, contrast, entropy, and angular second moment that are computed from the probability density function.

Difference Theorectic Features (DTF). In this feature extraction method, scale, illumination- invariant, and rotation based features were created from the correlated distributions of local and global intensity differences at different grey levels in the image. DTF computes eleven fetures based on the correlation values along with the horizonal, diagonal, and vertical directions to obtained the feature set.

Gabor Features. Gabor feature are used to compute the texture feature based on the wavelet function. The following wavelet function is used in this work [21]:

$$h(x, y) = \exp\left(-\alpha^{2\omega} \frac{x^2 + y^2}{2}\right) \exp(\omega\pi\alpha^\omega(x\cos\theta + y\sin\theta) \tag{2.5}$$

where $\alpha = \frac{1}{sqrt2}, \omega = 0, 1, 2, ...$ and $\theta \in [0, 2\pi]$.

Based on the wavelet function, The gabor features are computed at different frequencies, locations and orientations by convolving the input image with filters:

$$m(x, y) = L_h(I(x, y)) = |I(x, y) \times h(x, y)| \tag{2.6}$$

- *Color Features:* A space can be used to specify, create, and visualize color into distinct data representation [1]. The Uniformity in color on the surface of the apple makes it good for the open market. In this study, HSV color space is used to extract color features [13]. The HSV color space is close to how humans perceive color. It contains three elements as Hue (values lie between 0 to 360), Saturation, and View. where hue contains the color values, saturation specifies a range of gray values and view is the brightness of color. The formulation for the HSV features of a RGB image (I) as follow:

$$mx_{(i,j)} = max(I_{R(i,j)}, I_{G(i,j)}, I_{B(i,j)}); mn_{(i,j)} = min(I_{R(i,j)}, I_{G(i,j)}, I_{B(i,j)}) \tag{2.7}$$

$$H(i, j) = \begin{bmatrix} \frac{60*(I_{G(i,j)} - I_{B(i,j)})}{mx-mn} I_{R(i,j)} > max(I_{G(i,j)}, I_{B(i,j)}) \\ \frac{180*(I_{B(i,j)} - I_{R(i,j)})}{mx-mn} I_{G(i,j)} > max(I_{R(i,j)}, I_{B(i,j)}) \\ \frac{300*(I_{R(i,j)} - I_{G(i,j)})}{mx-mn} I_{B(i,j)} > max(I_{R(i,j)}, I_{G(i,j)}) \end{bmatrix} \tag{2.8}$$

- *Shape and Size Features:* The size and shape of the apples are important features due to their considerable ability to distinguish between different types of apples. In general, farmers typically look for shape and size features for the apple grading.

In this proposed approach, 18654 features are extracted from the segmented images of the apples.

2.4. Normalization. Normalization, a crucial preprocessing step that avoids larger numeric attribute values from influencing smaller numeric feature values, converts features into a common range. The main goal of the normalization techniques is to lessen the bias of the features that are numerically more significant than others. The performance of the z-score is superior for the feature selection process, hence it is employed in the proposed approach [37].

2.5. Hybrid Feature Selection and Feature Weighting. High-dimensional datasets might contain duplicate and irrelevant features. The effectiveness of the machine learning algorithms is influenced by the feature set's quality. Thus, optimal set of features are essential for improving the performance of machine learning algorithms. In this approach, A hybrid method has been proposed to select optimal features based on feature selection and feature weighting approach. First the redundant and irrelevant features are eliminated using the minimal-redundancy-maximum-relevance (mRMR) technique. The choosen features using the mRMR method tare both maximally different from one another and have the highest relevance to the target class [34]. After the selection of k number for features based of mRMR method, a feature weighting method based on the metaheuristic method has been used to assign weigth to the optimal selected features. The feature weighting helps in the selection of features based on the feature relevance. In this method, Sine Cosine based optimization algorithm has been utilized for the feature weighting of the apple dataset features. The deatail description is as follows:

2.5.1. Minimal-redundancy-maximum-relevance based feature selection. The feature selection algorithm is preresented by Ding and Peng and the minimum-redundancy maximum- relevance (mRMR) algorithm ranks a set of features by maximising their relevance and minimising their redundancy within the subset of features [9]. The goal is to reduce the duplication of a subset that has been chosen using a relevance metric. The F-test is used as a relevance measure in the first phase of the mRMR algorithm, and the Pearson's correlation between features is calculated as a redundancy measure. The remaining set of features is repeatedly chosen based on the mRMR score aluclated using the following formula:

$$score = max_{ie\Omega S} \left\{ F(i,s) - \left(\frac{1}{|S|} \sum_{j \in S} |c(i,j)| \right) \right\} \tag{2.9}$$

After the first feature, or the feature with the largest value of relevance (f-test ranked set), with the goal is chosen. F-test correlation difference (FCD) was the method employed in this investigation.After the selection of k number of feature usinh the mRMR feature selection method the selected feature then analysed using the feature weightning method for the selected of most relevant features.

2.5.2. Sine Cosine optimization algorithm based feature weightning. Sine cosine algorithm (SCA) is a population-based method for solving global optimisation issues [28]. SCA bases its search strategy on the properties of the sine and cosine trigonometric functions. SCA has demonstrated its effectiveness in both exploration and exploitation, and because of its prowess in exploration, it has been used to address a wide range of real-world problems. The approach uses the SCA algorithm as a weighting algorithm, giving features that are more important a larger weight than ones that are redundant or less relevant. In the SCA algorithm the position are updated based on the following formula after the random initialization of the position values:

$$P_i^{j+1} = P_i^j + rand_1 \times \sin(rand_2) \times |rand_3 \times D_i^j - P_i^j| \tag{2.10}$$

$$P_i^{j+1} = P_i^j + rand_1 \times \cos(rand_2) \times |rand_3 \times D_i^j - P_i^j| \tag{2.11}$$

where the current solution's position vector is represented by P_i^j at j^{th} iteration and i^{th} dimension. The $rand_1$, $rand_2$, and $rand_3$ denotes the random variables. The D_i^j shows the values of the position vector at the destination. The above equations are combined in the algorithm based on following formula:

$$P_i^{j+1} = \begin{cases} P_i^j + rand_1 \times \cos(rand_2) \times |rand_3 \times D_i^j - P_i^j|, rand_4 < 0.5, \\ P_i^j + rand_1 \times \cos(rand_2) \times |rand_3 \times D_i^j - P_i^j|, rand_4 \geq 0.5. \end{cases} \tag{2.12}$$

where $rand_4$ represents a random variable and its range in [0,1].

In the SCA algorithm, the initial features values are multiplied with weight values using a linear weight assignment technique [7] to increase accuracy. The following is the formulation of the linearly feature weighted data:

$$feat'_l = Weight_l \times feat_l \tag{2.13}$$

where $feat_l$ is the l^{th} feature and $Weight_l$ denotes the corresponding weight of l^{th} feature calcukated using SCA algorithm. By increasing the space of highly weighted features and decreasing the space of nominally weighted features, it modifies the feature space of the classification problem. Along with the selection of the relevant features using weighting method, the SCA method also selects the optimal parameteric values for the SVM classifier for the classification task.

2.6. Classification. SVM is a supervised learning model that analyses data used for regression and classification through associated learning computations. The main focus of the SVM classifier is on the hyper-plane from N-dimensional space that splits the data points associated with different classes [29]. In the proposed approach, the SVM classifier is utilized for the classification of apples. As the SVM looks for the most effective hyper-plane to use feature values to split the categories.

3. Results and Discussions. In this work, an approach for classification improvement of the apples is proposed. The experiments are performed on a Windows 10 Pro operating system running on the computer having the Intel® Xenon® CPU E5-2650 v3 (2.30 GHz) and 8 GB of RAM using the MATLAB 2019a platform. Ten-fold cross-validation is used for the classification performance evaluation.

In figure 3.1, classification accuracies have been achieved using the proposed approach for the different percentages of features selected using the mRMR approach illustrated.

It has been observed from figure 3.1 that the proposed approach achieves the highest accuracy is 92.27% at the 10% of features selected from the original set of apple features.

In figure 3.2, the accuracy comparision of the proposed approach with ReliefF [26], Mut-Inf [23], FSASL[10], zero-norm[46], Fisher[14], and Lasso [42] feature selection methods is illustrated.

Figure 3.2 indicates If we use all the features the accuracy was 87.94% but, the proposed approach achieves the highest 92.27% by selecting 10% of the features. This proves that the proposed approach is better in comparison to the other state-of-the-art methods as discussed in introduction section.

Friedman's mean rank test is used to verify the statistical significance of the proposed approach. In Table 3.1, Friedman's rank of the proposed approach in comparison to the well-regarded methods has been shown.

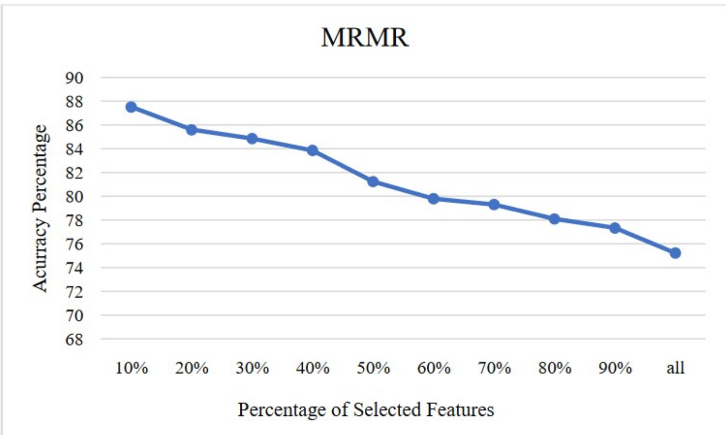

Fig. 3.1: Accuracy achieved using the proposed approach

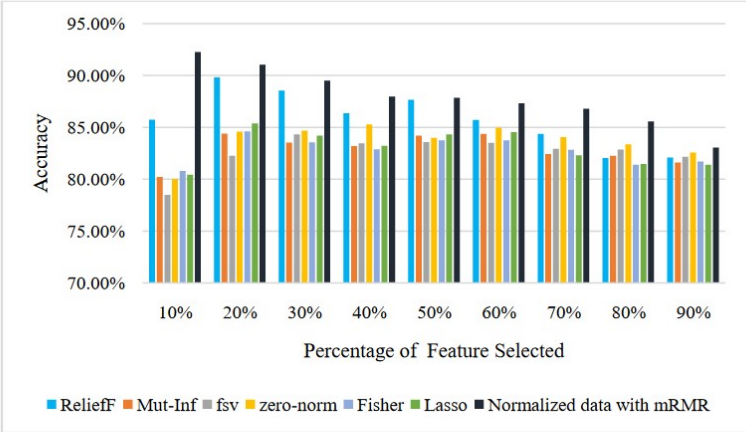

Fig. 3.2: Accuracy comparison of the proposed approach with state-of-the-art methods

Table 3.1: Friedman's mean rank of methods

Feature selection method	**Friedman Mean rank**
Normalized data with mRMR	1.11
Lasso	5.89
Fisher	6.33
Zero-Norm	3.67
Fsv	3.11
Mut-Inf	5.78
ReliefF	2.11

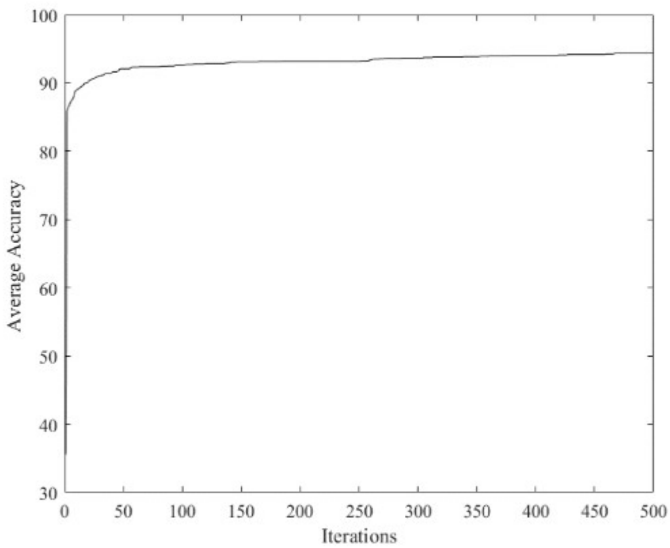

Fig. 3.3: Convergence curve of SCA based feature weighting

From Table 3.1, it has been observed that the proposed approach attains the lowest Friedman mean rank value with a p-value of 1.0825E-08 at a 5% of a confidence interval for the error values which shows the statistical significance of the normalized features selected using the mRMR approach. The seleceted 10% feature then fed into the SCA based feature weighting method and the convergence curve of attain using the algorithm is shown in the figure 3.3.

From the figure 3.3, it has been observed that the proposed SCA based feature weightning converge fater and attain accuracy of 94.3708% with the standard deviation value of 0.82 with optimized SVM classifier.

4. Conclusions. In this work, a methodology was proposed to enhance the classification of apples, structured into six phases. Initially, apple images were acquired, followed by segmentation to isolate regions. Features including texture, color, and shape & size were extracted and normalized to mitigate numerical biases. Redundant features were subsequently removed through a feature selection process. The final phase utilized a reduced feature set for classification. The method was validated using images of 158 royal delicious apples, showcasing different grades (A, B, C, D). SVM classifier is used for the classification of the apples into different classes. For the validation proposed approach is compared with the different existing state-of-the-art algorithms. The feature selection method attains the lowest Friedman's mean rank with the highest classification accuracy of 92.27% which proves the significance of using mRMR approach.The SCA based feature weighning approach attain accuracy of 94.3708% on the optimal set features obtained using the mRMR method. In the future, the proposed approach can be applied to the classification of the different fruits.

Data Availability. The data is made available on demand.

REFERENCES

[1] H. Afrisal, M. Faris, G. Utomo P., L. Grezelda, I. Soesanti, and M. Andri F. Portable smart sorting and grading machine for fruits using computer vision. In *Proceeding - 2013 Int. Conf. Comput. Control. Informatics Its Appl. "Recent Challenges Comput. Control Informatics", IC3INA 2013*, pages 71–75, 2013.

[2] M. A. H. Ali and K. W. Thai. Automated fruit grading system. In *2017 IEEE 3rd International Symposium in Robotics and Manufacturing Automation (ROMA)*, pages 1–6, 2017.

[3] S. K. Behera, A. K. Rath, and P. K. Sethy. Maturity status classification of papaya fruits based on machine learning and transfer learning approach. *Inf. Process. Agric.*, 8(2):244–250, Jun. 2021.

[4] A. Bhargava and A. Bansal. Automatic detection and grading of multiple fruits by machine learning. *Food Anal. Methods*, 13(3):751–761, Mar. 2020.

[5] A. K. Bhatt, D. K. Ghosh, S. J. Darokar, J. S. Sikarwar, and P. R. Kadu. Analysis of volatile compounds in ripening of apple (malus domestica) fruits using gas chromatography–mass spectrometry (gc–ms). *Curr. Sci.*, 115(5):891–895, 2018.

[6] M.H. Dairath, M.W. Akram, M.A. Mehmood, H.U. Sarwar, M.Z. Akram, M.M. Omar, and M. Faheem. Computer vision-based prototype robotic picking cum grading system for fruits. *Smart Agric. Technol.*, 4:100210, 2023.

[7] S. Dalwinder, S. Birmohan, and K. Manpreet. Simultaneous feature weighting and parameter determination of neural networks using ant lion optimization for the classification of breast cancer. *Biocybern. Biomed. Eng.*, 40(1):337–351, Jan. 2020.

[8] F. A. Diaz-Gonzalez, J. Vuelvas, C. A. Correa, V. E. Vallejo, and D. Patino. Machine learning and remote sensing techniques applied to estimate soil indicators – review. *Ecol. Indic.*, 135:108517, Feb. 2022.

[9] C. Ding and H. Peng. Minimum redundancy feature selection from microarray gene expression data. In *Computational Systems Bioinformatics. CSB2003. Procs of 2003 IEEE Bioinformatics Conference*, 523–528. IEEE Comp.Soc,2003.

[10] L. Du and Y. D. Shen. Unsupervised feature selection with adaptive structure learning. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 209–218, New York, NY, USA, Aug. 2015. ACM.

[11] S. D. Golombek and M. M. Blanke. Orchard management strategies to reduce bruises on apples in india: a review. *Vegetos*, 35(1):1–8, Mar. 2022.

[12] J. F. S. Gomes and F. R. Leta. Applications of computer vision techniques in the agriculture and food industry: a review. *Eur. Food Res. Technol.*, 235(6):989–1000, 2012.

[13] R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital Image Processing Using MATLAB: Pearson Prentice Hall*. Pearson Prentice Hall, Up. Saddle River, New Jersey, 2004.

[14] Q. Gu, Z. Li, and J. Han. Generalized fisher score for feature selection. In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence, UAI 2011*, pages 266–273. AUAI Press, Feb. 2011.

[15] H. Guo, Y. Tan, and W. Li. Surface texture detection of double-feature apple based on computer vision. In *Computational Intelligence in Robotics and Automation*, pages 117–127. 2014.

[16] M. Halstead, C. Mccool, S. Denman, T. Perez, and C. Fookes. Fruit quantity and quality estimation using a robotic vision system. 2018.

[17] R. M. Haralick, K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.*, (6):610–621, 1973.

[18] Jin Huang and C. X. Ling. Using auc and accuracy in evaluating learning algorithms. *IEEE Trans. Knowl. Data Eng.*, 17(3):299–310, Mar. 2005.

[19] K. Jain and N. Choudhary. Comparative analysis of machine learning techniques for predicting production capability of crop yield. *Int. J. Syst. Assur. Eng. Manag.*, 13(S1):583–593, Mar. 2022.

[20] P. Jolly and S. Raman. Analyzing surface defects in apples using gabor features. In *Proceedings - 12th International Conference on Signal Image Technology and Internet-Based Systems, SITIS 2016*, pages 178–185. Institute of Electrical and Electronics Engineers Inc., Apr. 2017.

[21] Z. Juan and C. Xiao-Ping. Field pest identification by an improved gabor texture segmentation scheme. *New Zeal. J. Agric. Res.*, 50(5):719–723, 2007.

[22] K. Kira and L. A. Rendell. A practical approach to feature selection. In *Machine Learning Proceedings 1992*, pages 249–256. Elsevier, 1992.

[23] A. Kraskov, H. Stögbauer, and P. Grassberger. Estimating mutual information. *Phys. Rev. E*, 69(6):066138, Jun. 2004.

[24] D. J. Lee, J. K. Archibald, and G. Xiong. Rapid color grading for fruit quality evaluation using direct color mapping. *IEEE Trans. Autom. Sci. Eng.*, 8(2):292–302, Apr. 2011.

[25] Q. Li, M. Wang, and W. Gu. Computer vision based system for apple surface defect detection. *Comput. Electron. Agric.*, 36(2–3):215–223, Nov. 2002.

[26] H. Liu and H. Motoda. *Computational Methods of Feature Selection.* Chapman & Hall/CRC, 2007.

[27] T. Y. Melesse, M. Bollo, V. Di Pasquale, F. Centro, and S. Riemma. Machine learning-based digital twin for monitoring fruit quality evolution. *Procedia Comput. Sci.*, 200:13–20, 2022.

[28] S. Mirjalili. Sca: A sine cosine algorithm for solving optimization problems. *Knowledge-Based Syst.*, 96:120–133, Mar. 2016.

[29] A. Mizushima and R. Lu. An image segmentation method for apple sorting and grading using support vector machine and otsu's method. *Comput. Electron. Agric.*, 94:29–37, Jun. 2013.

[30] P. Moallem, A. Serajoddin, and H. Pourghassem. Computer vision-based apple grading for golden delicious apples based on surface features. *Inf. Process. Agric.*, 4(1):33–40, 2017.

[31] G. Muhammad. Date fruits classification using texture descriptors&shape-size features. *Eng.Appl.Artif.Intell.*,37:361–36,2015.

[32] S. Naik and B. Patel. Machine vision based fruit classification and grading - a review. *Int.J.Comput.Appl.*,170(9):22–34,2017.

[33] J. F. I. Nturambirwe and U. L. Opara. Machine learning applications to non-destructive defect detection in horticultural products. *Biosyst. Eng.*, 189:60–83, Jan. 2020.

[34] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(8):1226–1238, Aug. 2005.

[35] B. T. Pham, L. H. Son, T. A. Hoang, D. M. Nguyen, and D. Tien Bui. Prediction of shear strength of soft soil using machine learning methods. *CATENA*, 166:181–191, Jul. 2018.

[36] A. Sengupta, A. Mukherjee, A. Das, and D. De. Agristick: An iot-enabled agricultural appliance to measure growth of jackfruit using 2-axis joystick. *IEEE Instrum. Meas. Mag.*, 25(3):58–62, May 2022.

[37] D. Singh and B. Singh. Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.*, 97:105524, Dec. 2020.

[38] M. M. Sofu, O. Er, M. C. Kayacan, and B. Cetişli. Design of an automatic apple sorting system using machine vision. *Comput. Electron. Agric.*, 127:395–405, 2016.

[39] M. Suresha, N. A. Shilpa, and B. Soumya. Apples grading based on svm classifier. *Int. J. Comput. Appl.*, 975:8878, 2012.

[40] S. Susan and M. Hanmandlu. Difference theoretic feature set for scale-, illumination- and rotation-invariant texture classification. *IET Image Process.*, 7(8):725–732, Nov. 2013.

[41] H. A. Le Thi, V. V. Nguyen, and S. Ouchani. Gene selection for cancer classification using dca. In *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, volume 5139 LNAI, pages 62–72, 2008.

[42] R. Tibshirani. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B*, 58(1):267–288, Jan. 1996.

[43] M. Tripathi. Analysis of convolutional neural network based image classification techniques. *J. Innov. Image Process.*, 3:100–117, 2021.

[44] S. Veerashetty and N. B. Patil. Novel lbp based texture descriptor for rotation, illumination and scale invariance for image texture analysis and classification using multi-kernel svm. *Multimed. Tools Appl.*, 79(15–16):9935–9955, Apr. 2020.

[45] T. Vijayakumar. Posed inverse problem rectification using novel deep convolutional neural network. *J. Innov. Image Process.*, 2:121–127, 2020.

[46] J. Weston, A. Elisseeff, B. Schölkopf, and M. Tipping. Use of the zero-norm with linear models and kernel methods. *J. Mach. Learn. Res.*, 3:1439–1461, 2003.

[47] J. S. Weszka, C. R. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Trans. Syst. Man Cybern.*, SMC-6(4):269–285, 1976.

[48] M. Yang, P. Kumar, J. Bhola, and M. Shabaz. Development of image recognition software based on artificial intelligence algorithm for the efficient sorting of apple fruit. *Int. J. Syst. Assur. Eng. Manag.*, 13(S1):322–330, Mar. 2022.

[49] Jinli Yuan, Zhitao Guo, and Dawei Yue. The apple color grading based on pso and svm. In *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, pages 5198–5201. IEEE, Aug. 2011.

Edited by: Manish Gupta

Special issue on: Recent Advancements in Machine Intelligence and Smart Systems

Received: May 13, 2024

Accepted: Jun 11, 2024