Quality of performance evaluation of ten machine learning algorithms in classifying thirteen types of apple fruits

Nashaat M. Hussain Hassan[1,2], Basma Ramadan Gamal Elshoky[3,4], A. M. Mabrouk[1]
[1]Faculty of Engineering and Technology, Badr University in Cairo (BUC), Badr, Egypt
[2]Department of Electronics and Electrical Communication Faculty of Engineering, Fayoum University, Fayoum, Egypt
[3]Departement of information Technology, Korean Egyptian Faculty for Industry and Energy Technology, Beni Suef Technological University, Beni Suef, Egypt
[4]Computer Science, Faculty of Science, Minia University, Minia, Egypt

Article Info

Article history:

Received Oct 4, 2022
Revised Nov 25, 2022
Accepted Nov 29, 2022

Keywords:

Classification
Features extractions
First apple fruits
Machine learning algorithms
Processing speed
Quality of performance

ABSTRACT

Recently, computer vision technology has become essential for the automatic, accurate, and fast classification of fruits. Actually, there are many challenges in separating the types of fruits that are somewhat similar, such as apples, pears, and peaches. However, the challenges become more difficult if the separation is on different varieties of the same fruit. While the difficulty doubles if the classification takes place with a large number of different varieties of the same fruit. Most of the literature which is presented in this regard, and which is relied on the use of machine learning techniques lacked the following: first; the focus was on certain technologies such as k-nearest neighbor (KNN), support vector machine (SVM) without looking at many other machine learning techniques. Second; the literature was concerned only with measuring the accuracy of the techniques that are used, without looking at the relationship between the accuracy and processing speed (computation times). This manuscript aims to study and analyze the results of measuring accuracy and computation times for ten machine-learning techniques in order to identify and classify thirteen types of apples. After studying and analyzing the results, many observations were made, which will be referred to in the results section.

Corresponding Author:

Nashaat M. Hussain Hassan
Department of Electronics and Electrical Communication Faculty of Engineering, Fayoum University
Qesm Al Fayoum, Faiyum, Faiyum Governorate 2933110, Egypt
Email: nmh01@fayoum.edu.eg

1. INTRODUCTION

To assure the production, goodness, thoroughness, competitively, and cost reduction of apple fruitiness products, an automatic classification of apple fruits [1]–[5] has become essential for companies selling, manufacturing and exporting apples. Actually, the automatic classification of a large number of different types of apples with the best quality and at the highest speed is not easy at all. Many methods were developed to classify apple fruitiness depending on the shape, color, texture, and appearance features [6]–[9]. An automatic classification always works depending on the application of computer vision techniques [10]–[12]. Machine learning techniques [13]–[16] are a part of computer vision techniques. To the fact that the great overlap between the different types of apples, in colors, shapes, sizes, and textures, increases the difficulty of classification for these types. This makes traditional classification techniques unsuitable for executing processors with such difficulty and for this amount of data, and at high speeds. Therefore, machine learning technologies have become one of the most powerful solutions for this type of process. In my opinion, important aspects were missing from much of the literature that dealt with the use of machine learning techniques to classify the different types of fruits and vegetables. These aspects include the following: i) Much of this literature has focused on the use and development of support vector machine (SVM) [17], and k-nearest neighbor (KNN) [17] technologies, without concern for the use or the development of other machine learning techniques such as LR, LDA, CART, NB, AB, GBM, RF, and ET [18]. ii) Second; the literature was concerned only with measuring the accuracy of the techniques that are used, without looking at the relationship between the accuracy and processing speed (computation times). In the following, a review of samples of the literature contains what has been referred to. In the research paper entitled "Classification of Some Fruits Using Image Processing and Machine Learning", which is authored by Koç and Vatandaş [19], the author presented the results of classification accuracy for five techniques KNN, decision tree (DT), Naive Bayes classification and multilayer perceptron neural network (MLP) [19] algorithms were used in order to classify 300 types of fruits. Although the author tested five techniques and this is good, but, he did not mention anything about the results of the processing speed of these techniques, which is necessary to refer to as long as it is related to techniques used in computer vision applications. In the research paper entitled "Machine Learning for Apple Fruit Diseases Classification System", which is for the author El-aziz *et al.* [20]. It is noted that the author focused on a comparison review between the results of the classification accuracy for two machine learning techniques, namely SVM and KNN, before and after modification, in order to separate four types of apple diseases. So, the focus of this paper is on presenting classification accuracy results for only two machine learning techniques (SVM, and KNN), in addition to not referring to processing speed results for these techniques.

Also, the research paper, which is entitled different apple varieties classification using KNN and MLP algorithms, which is by the Sabancı [21], the research also dealt with a comparison between the results of the classification accuracy for only two machine learning techniques, namely KNN and MLP, in order to separate three types of apples without looking at the results of processing speed as well. So, it is noted from the above that much of the literature depends on the use and development, testing the classification accuracy of these techniques on a limited number of varieties, in addition to not looking at the results of the processing speed of these techniques. Therefore, this manuscript aims to study and analyze the results of measuring accuracy and processing speed for ten machine learning techniques (LR, LDA, CART, NB, KNN, SVM, AB, GBM, RF, and ET) in order to identify and classify thirteen types of apples. In order to ensure the accuracy of the results obtained, the ten techniques were tested on a collection of 6,404 images. In order to ensure the accuracy of the results obtained, the ten techniques were tested on a collection of 6,404 images. After studying and analyzing the results of testing the ten techniques on this quantity of apples, the following was noted: first; It is not only techniques: SVM: 97%, and KNN: 96% that achieve high accuracy, but there are other techniques that achieve high accuracy as well, such as ET: 96%, RF: 95%, GBM: 95%, LR: 94%. Second; SVM and KNN technologies are not the fastest in processing (53.62 s, 63.86 s), but there are other technologies such as ET, and RF that offer the same accuracy as SVM and KNN, but with a much faster processing speed (3.65 s, 10.94 s) up to 20 times the speed of SVM, and KNN. Third; NB technology is the fastest ever, as it achieved an efficiency of 84% in a time of 2.14 s. While the slowest technology at all is gradient boosting machine (GBM), which achieved an efficiency of 95% in a time of 6066.14 s. Forth; The highest accuracy which is achieved with the lowest time is for the ET algorithm (96% at 3.65 s) also the RF algorithm achieved a good accuracy and a good time (95% and 10.94 s). While the AB algorithm achieved a very low accuracy of 34% in time 112.47 s.

There are many other important notes on the results that will be presented in the section explaining the results. This paper is constructed as: in the second section, an overview of the machine learning algorithm, results, and discussion are demonstrated in section three. In the last section, the conclusion and future lines are presented.

2. METHOD

Our concern in this work is related to the evaluation of the performance quality and processing speed of the ten most important machine learning techniques, in order to classify different types of apples. Important things that were required to be known in this research, for example: i) What are the most appropriate techniques among the ten techniques for grading apples? ii) What are the technologies that have the lowest performance quality among the ten technologies that are affected greatly, the greater the variety of apples to be classified? iii) What are the technologies that perform with the best efficiency and the fastest processing speed? iv) What are the slowest technologies ever? And v) A lot of things like that. So, the calculation methods and a brief discussion of the ten machine learning algorithms, which were used in the three different types of apple fruit classification tasks, will be introduced in this section.

2.1. Decision tree

Classification and regression tree (CART) [22] is constructed on the decision tree approach by combination regression and classification trees. It split the data recursively then fits a prediction model for every partition. CART algorithm has advantages: it is nonparametric, flexible, can adjust in time, no assumptions, and computationally fast. Giniplitting rule approach is used with a classification tree for splitting data into right and left smaller parts. Then it searches among learning sample data to find the best class and insulate it from the remnant. Gini splitting math is in (1), where:

- The parent node t_{parent} is split into P_{left} and P_{right}, n_{left} and n_{right} are child node of P_{left} and $P_{right.}$
- P(c|n) refers to the conditioned probability for class C that belongs to node n.
- Y refers to a variable matrix with Z number of variables $y_{s.}$
- $Y^r{}_s$ refers to the best value for splitting variable $Y_{s.}$
- Crefers to an index of the class.

$$arg(max)\text{Ys} <= \text{Yrs}, s = 1, \dots, Z \ [-\sum_{c=1}^{c} p^2(c|n_p)+p_l\sum_{c=1}^{c} p^2(c|n_l)+p_r\sum_{c=1}^{c} p^2(c|n_r). \quad (1)$$

2.2. Linear discriminant analysis

A probability method used to dimensionality reduction also data classification [23]. The linear discriminant function of the ℓ th group in the (2), where:

- X is the value that will classify
- Symbol μ (multidimensional mean vector) is a parameter
- C refers to no. of classes
- P_r refers to prior probability

$$x`\Sigma^{-1}\mu\ell - 0.5\mu'\ell\ \Sigma^{-1}\ \mu\ell + \log(P_r[Y = C\ell]). \quad (2)$$

2.3. Ada Boost

Boosting is a machine learning technique that gathers several close weak and inexact rules for creating a highly perfect prediction rule [24]. The primary concept of boosting means that insert new models into the ensemble sequent [24]. Ada Boost (AB) is an ensemble boosting algorithm. Boosting is based on combine several minimal performing classifiers in an interacting way to produce a preferable performing classifier. The concept of ada-boost is to set the weights of classifiers and the training data sample in each repetition such that it assures valid predictions of uncommon instances [24]. The formula of Ada Boost function shown (3) [24].

$$F(x) = \sum_{i=1}^{i} \alpha_i h_i(x) \quad (3)$$

where:

- h_i weak hypotheses
- α_i the weight
- i the number of round
- x is the data

2.4. Gradient boosting machine

The notion of GBM [25] is constructing the new base-learners to be utmost correlated with the passive gradient of the loss function and associated with the all ensemble. Specify both the loss function Ψ (y_predict, function) and the base-learner l (z, θ) models on demand. Mathematical equation of GBM is shown as:

$$(\rho_t, \theta_t) = \arg\min_{\rho,\theta} \sum_{i=1}^{N} [-g_t(w_i) + \rho h\ (w_i, \theta)]^2 \quad (4)$$

where:

- ρ is the estimate function
- $g_t(w_i)$ negative gradient
- h (w_i, θ) custom base-learner
- w_i refers to the input variable
- θ is a parameter

2.5. Support vector machine

SVM is an universal learning machine parameterized by set weights and support vectors to make the decision, also characterized by a kernel function [26], [27]. The maximum margin classifier produces a D(x), that is maximize the stop between the border and the data [26], [27]. The mathematical constructs of SVM:

$$D(u) = \beta_0 + \sum_{j=1}^{p} \beta_j u_j = \beta_0 + \sum_{i=1}^{n} y_i a_i x'_i u \quad (5)$$

where:
- D(u) a decision function for new sample u
- β is a parameter
- α_i the parameter estimates
- x and y are the data

2.6. Random forest

The random forest is an ensemble algorithm that utilizes rankers based on bagging and specimen features [28]. Bagging refers to the proceeding of collecting various decision trees then calculating the average of them. General basic steps of random forest algorithm are shown in the following:
- Step 1: First, the algorithm starts by selecting the random samples from the selected dataset.
- Step 2: Next, for every sample, a decision tree will be constructed in this algorithm. Then, results of prediction from every decision tree will be extracted.
- Step 3: For every predicted result a voting will be accomplished, at this step.
- Step 4: At the end, selecting the most supported prediction outcome as the last prediction outcome.

2.7. Naive Bayes

As a simple learning method that used the Bayes rule. It is used information in-sample data to calculating the posterior probability P (y | x) [29]. It can use for binary (two-class) and multi-class classification. The Naive Bayes model simplifies the probabilities of the predictor values by assuming that all of the predictors are independent of the others [29]. The probability densities for every predictor are in the following mathematical in (6).

$$P_r[X|Y = C\ell] = \Pi^{p}_{j=i} P_r[X_j|Y = C\ell] \quad (6)$$

where:
- P_r refers to the prior probability for the outcome
- C refers to no. of classes
- X is an input data
- Y refers to an output class

2.8. Extra trees

Extra trees (ET) used the classical top-down procedure to build an ensemble of decision trees. It splits nodes by choosing cut-points fully at random and uses the whole learning sample to grow the trees [30]. The basic steps of ET algorithm are shown in the following steps:
- Step 1: Bring in the requested libraries
- Step 2: Reading and spring-cleaning the dataset
- Step 3: Structuring the extra trees forest and calculating the importance discrete feature
- Step 4: Imagining and matching the results

2.9. K-neighbors

KNN used the K-closest to learn from the training part samples for predicting the class of a new sample. The K-closest learned from samples via the distance metric like Demidova [31]. Minkowski distance is a generalization of Euclidean distance and is defined in (7).

$$\text{Minkowski distance} = \left(\sum_{j=i}^{p} |z_{aj} - z_{bj}|^{q}\right)^{\frac{1}{q}} \quad (7)$$

Where: z_a and z_b refer to as two samples.

2.10. Logistic regression

LR [32] is a mathematical modeling approach that can be used to describe the relationship of several X_s to a dichotomous dependent variable, such as D the logistic model is popular because the logistic function, on which the model is working based on the following steps:

- Step 1: Pre-processing of the dataset is done
- Step 2: Appropriate logistic regression to the training process is done
- Step 3: Expecting of the test results is done at this step
- Step 4: Testing the accuracy of the obtained results is accomplished is realized at this step
- Step 5: at the end, visualizing of the results of the tested dataset is implemented

Calculation of the LR algorithm is presented in (8), (9) as the following:

$$\log\left(\frac{p}{1-p}\right) = \beta_{0+} \beta_1 x_{1+\ldots+} \beta_p x_p \tag{8}$$

$$p = \frac{1}{1+\exp\left[-(\beta_0+\beta_1 x_1+\cdots+\beta p X p)\right]} \tag{9}$$

where: P is the number of predictors. Symbol β refer to the parameter.

3. RESULTS AND DISCUSSION

The dataset which is used in this work is called Fruits-360. It is found in the reference [33]. A collection of 6,404 images for 13 apple fruits classes were used in the classification train process and 2134 for the test. A software implementation of three classifications of different types of Apple fruits (by applying ten algorithms of machine learning algorithms) was implemented. The first type is to classify between four types of apple fruits, the second type is to classify between nine types, and the last type is to classify thirteen types of apple fruits. In the first one, collections of 1918 images for four apple fruits classes were used in the classification train process. In the second one, Collections of 4,486 images for nine apple fruits classes were used in the training process, and 1,492 images for the test. In the last one, a collection of 6,404 images for 13 apple fruits classes were used in the classification train process and 2,134 for the test.

Our concern in this research paper is concentrated on monitoring the rapid and accurate one of the ten different techniques of machine learning technologies (LR, LDA, CART, NB, KNN, SVM, AB, GBM, RF, and ET) for classifying thirteen various types of apple fruitiness. Samples of the results have shown in Tables 1, 2, and 3. After many classification processes with different cases of apple collections, the obtained results indicate the following: The SVM technology achieved 97% (average quality of performance over the three types of classifications which are made on apples) in 53.63s; KNN achieved 96% in 63.86s, while ET achieved 96% in only 3.55 second (ET computation times less than the computation times of KNN and SVM by 20 times). The RF technology achieved 95% in 10.94 seconds (good accuracy at low computation time). It is also noted that the technology GBM achieved 95% in 6066.14 seconds (good accuracy but a very long time for computations). It is also noted that the technology AB achieved 34% Accuracy in 112.47 seconds (very bad accuracy and so bad computation times).

Another side can be observed from the obtained results, which is, for the first type of classification (four types of apple fruits), eight algorithms (from the ten ML algorithms) achieved an accuracy of more than 97 %. (KNN, SVM, ET, RF, CART, LDA, and LR), while AB achieved the lowest accuracy 58%. In the second type of classification (nine types of apple fruits) six algorithms achieved an accuracy of more than 94% (RF: 94%, LR: 95%, ET: 96%, GBM: 96%, SVM: 98%, KNN: 98%), and stay the AB algorithm giving the lowest accuracy: 27%, at the last type of classification (thirteen types of apple fruits), five algorithms from the ten achieved accuracy equal to 92% (ET: 92%, RF: 92%, GBM: 92, KNN: 92%, and SVM: 94%). On the other side, it is observed that for the first type (four types of apple fruits) only four algorithms were achieved in computation time of fewer than 5 seconds (ET:1.14s to achieve 100%, RF: 2.95 s to achieve 100%, NBs: 0.87 s to achieve 90%, CART:3.27 s to achieve 97%). In the second type of classification (nine apple fruits) only four algorithms were achieved in computation times less than 15 seconds (ET: 3.62 s to achieve 96% accuracy, RF: 11.64 s to achieve 94%, NB: 2.11 seconds to achieve 83%, and CART: 14.62 to achieve 80%). In the last type of classification (thirteen types of apple fruits), only four algorithms were achieved in time less than 25 seconds (ET: 5.86 s to achieve 92% accuracy, RF: 18.26 s to achieve 92%, NB: 3.45 seconds to achieve 79%, and CART: 23.19 s to achieve 79%).

As a conclusion of these all results, we can conclude the following: first, the biggest accuracy which was achieved with the lowest time is for the ET alalgorithm (96% and 3.65 s). Second, the RF algorithm achieved good accuracy in good time (95% and 10.94 s). Third, the SVM achieved accuracy at 97% but in 53.62 s, and the KNN algorithm achieved accuracy at 96% in 63.86 s (long time). Fourth, it is observed also that the NB's algorithm is a very fast algorithm, were achieved an accuracy of 84% in only 2.14 s. Fifth, the AB algorithm achieved a very low accuracy of 34% in a very long time 112.47 s.

Table 1. Results of quality of performance of the ten ML algorithms over the three cases of classification

Classification type	LR	LDA	CART	NB	KNN	SVM	AB	GBM	RF	ET
Four types	1.00	0.99	0.97	0.90	1.00	1.00	0.58	0.99	1.00	1.00
Nine types	0.95	0.88	0.80	0.83	0.98	0.98	0.27	0.96	0.94	0.96
Thirteen types	0.86	0.84	0.79	0.79	0.92	0.94	0.19	0.92	0.92	0.92
Average	0.94	0.90	0.85	0.84	0.96	0.97	0.34	0.95	0.95	0.96

Table 2. Results of computation times in seconds of the ten ML algorithms over the three cases of classification

Classification types	LR	LDA	CART	NB	KNN	SVM	AB	GBM	RF	ET
Four types	28.07	13.16	3.27	0.87	16.11	5.32	52.35	1062.53	2.95	1.14
Nine types	110.33	70.7	14.62	2.11	57.29	47.85	118.38	5585	11.61	3.62
Thirteen types	223	191.5	23.19	3.45	118.2	107.7	166.7	11550.9	18.26	5.89
Average	120.46	91.78	13.69	2.14	63.86	53.62	112.47	6066.14	10.94	3.55

Table 3. Results of computation time's average and quality of performance average of the ten ML algorithms over the three types of classification

Evaluation measures	LR	LDA	CART	NB	KNN	SVM	AB	GBM	RF
Computation times	120.46	91.78	13.69	2.14	63.86	53.62	112.47	6066.14	10.94
Average	0.94	0.90	0.85	0.84	0.96	0.97	0.34	0.95	0.95

4. CONCLUSION

Software implementation of ten machine learning algorithms to classify a large number of different types of apple fruits in order to study and analyze the accuracy and computation time of that algorithms are realized in the presented research paper. After many classification processes with different cases of apple collections, many observations were made, which will be referred to in the results section. Future work will focus on the following: first, applying deep learning algorithms to compare the results of ten machine learning algorithms (which were tested in this research paper) and the results of testing the deep learning algorithms (for accuracy and computation times) to find out the best techniques for classifying different types of apple fruits. Second, think about the hardware implementation of the technique that gives the best results (high accuracy in a very low time), so that it can be used in industrial enterprises that rely on computer vision techniques.

REFERENCES

[1] F. Femling, A. Olsson, and F. Alonso-Fernandez, “Vegetable identification using machine learning for retail applications,” in *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, Nov. 2018, pp. 9–15. doi: 10.1109/SITIS.2018.00013.

[2] A. A. Nashat and N. M. H. Hassan, “Automatic segmentation and classification of olive fruits batches based on discrete wavelet transform and visual perceptual texture features,” *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 16, no. 01, p. 1850003, Jan. 2018, doi: 10.1142/S0219691318500030.

[3] N. M. H. Hassan and A. A. Nashat, “New effective techniques for automatic detection and classification of external olive fruits defects based on image processing techniques,” *Multidimens. Syst. Signal Process.*, vol. 30, no. 2, pp. 571–589, Apr. 2019, doi: 10.1007/s11045-018-0573-5.

[4] F. M. A. Mazen and A. A. Nashat, “Ripeness classification of bananas using an artificial neural network,” *Arab. J. Sci. Eng.*, vol. 44, no. 8, pp. 6901–6910, Aug. 2019, doi: 10.1007/s13369-018-03695-5.

[5] K. Tan, W. S. Lee, H. Gan, and S. Wang, “Recognising blueberry fruit of different maturity using histogram oriented gradients and colour features in outdoor scenes,” *Biosyst. Eng.*, vol. 176, pp. 59–72, Dec. 2018, doi: 10.1016/j.biosystemseng.2018.08.011.

[6] S. Sabzi, Y. Abbaspour-Gilandeh, and G. García-Mateos, “A new approach for visual identification of orange varieties using neural networks and metaheuristic algorithms,” *Inf. Process. Agric.*, vol. 5, no. 1, pp. 162–172, Mar. 2018, doi: 10.1016/j.inpa.2017.09.002.

[7] A. Pande, M. Munot, R. Sreeemathy, and R. V. Bakare, “An efficient approach to fruit classification and grading using deep convolutional neural network,” in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, Mar. 2019, pp. 1–7. doi: 10.1109/I2CT45611.2019.9033957.

[8] J. L. Rojas-Aranda, J. I. Nunez-Varela, J. C. Cuevas-Tello, and G. Rangel-Ramirez, “Fruit classification for retail stores using deep learning,” 2020, pp. 3–13. doi: 10.1007/978-3-030-49076-8-1.

[9] M. S. Hossain, M. Al-Hammadi, and G. Muhammad, “Automatic fruit classification using deep learning for industrial applications,” *IEEE Trans. Ind. Informatics*, vol. 15, no. 2, pp. 1027–1034, Feb. 2019, doi: 10.1109/TII.2018.2875149.

[10] R. L. Sekar, N. Ambika, V. Divya, and T. Kowsalya, “Fruit classification system using computer vision : a review,” *Int. J. Trend Res. Dev.*, vol. 5, no. 1, pp. 22–26, 2018.

[11] N. Ismail and O. A. Malik, “Real-time visual inspection system for grading fruits using computer vision and deep learning techniques,” *Inf. Process. Agric.*, vol. 9, no. 1, pp. 24–37, Mar. 2022, doi: 10.1016/j.inpa.2021.01.005.

[12] S. Naik and B. Patel, “Machine vision based fruit classification and grading - a review,” *Int. J. Comput. Appl.*, vol. 170, no. 9, pp. 22–34, Jul. 2017, doi: 10.5120/ijca2017914937.

[13] A. Kader, S. Sharif, P. Bhowmick, F. H. Mim, and A. Y. Srizon, “Effective workflow for high-performance recognition of fruits using machine learning approaches,” *Int. Res. J. Eng. Technol.*, vol. 07, no. 02, 2020.

[14] P. Moallem, A. Serajoddin, and H. Pourghassem, “Computer vision-based apple grading for golden delicious apples based on surface features,” *Inf. Process. Agric.*, vol. 4, no. 1, pp. 33–40, Mar. 2017, doi: 10.1016/j.inpa.2016.10.003.

[15] L. S. P. Annabel, T. Annapoorani, and P. Deepalakshmi, “Machine learning for plant leaf disease detection and classification – a review,” in *2019 International Conference on Communication and Signal Processing (ICCSP)*, Apr. 2019, pp. 0538–0542. doi: 10.1109/ICCSP.2019.8698004.

[16] H. Mureşan and M. Oltean, “Fruit recognition from images using deep learning,” *Acta Univ. Sapientiae, Inform.*, vol. 10, no. 1, pp. 26–42, Aug. 2018, doi: 10.2478/ausi-2018-0002.

[17] A. Nosseir and S. E. A. Ahmed, “Automatic classification for fruits’ types and identification of rotten ones using k-NN and SVM,” *Int. J. Online Biomed. Eng.*, vol. 15, no. 03, p. 47, Feb. 2019, doi: 10.3991/ijoe.v15i03.9832.

[18] H. M. W. M. Hippola, D. P. WaduMesthri, R. M. T. P. Rajakaruna, L. Yasakethu, and M. Rajapaksha, “Machine learning based classification of ripening and decay stages of Mango (Mangifera indica L.) cv. Tom EJC,” in *2022 2nd International Conference on Image Processing and Robotics (ICIPRob)*, Mar. 2022, pp. 1–6. doi: 10.1109/ICIPRob54042.2022.9798722.

[19] D. G. Koç and M. Vatandaş, “Classification of some fruits using image processing and machine learning,” *Turkish J. Agric. - Food Sci. Technol.*, vol. 9, no. 12, pp. 2189–2196, Dec. 2021, doi: 10.24925/turjaf.v9i12.2189-2196.4445.

[20] A. A. A. El-aziz, A. Darwish, D. Oliva, and A. E. Hassanien, “Machine learning for apple fruit diseases classification system,” 2020, pp. 16–25. doi: 10.1007/978-3-030-44289-7_2.

[21] K. Sabancı, “Different apple varieties classification using knn and mlp algorithms,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 4, no. Special Issue-1, pp. 166–169, Dec. 2016, doi: 10.18201/ijisae.2016SpecialIssue-146967.

[22] N. Chaithra, B. Madhu, K. Umamaheswari, and S. Balasubramanian, “Classification and regression tree (CART) algorithm for the prediction of ischemic heart disease,” *Int. J. Futur. Gener. Commun. Netw.*, vol. 12, no. 4, pp. 196–206, 2020.

[23] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, “Linear discriminant analysis: A detailed tutorial,” *AI Commun.*, vol. 30, no. 2, pp. 169–190, May 2017, doi: 10.3233/AIC-170729.

[24] S. Wang and Q. Chen, “The study of multiple classes boosting classification method based on local similarity,” *Algorithms*, vol. 14, no. 2, p. 37, Jan. 2021, doi: 10.3390/a14020037.

[25] D. A. McCarty, H. W. Kim, and H. K. Lee, “Evaluation of light gradient boosted machine learning technique in large scale land use and land cover classification,” *Environments*, vol. 7, no. 10, p. 84, Oct. 2020, doi: 10.3390/environments7100084.

[26] F. R. Lumbanraja, E. Fitri, Ardiansyah, A. Junaidi, and R. Prabowo, “Abstract Classification Using Support Vector Machine Algorithm (Case Study: Abstract in a Computer Science Journal),” *J. Phys. Conf. Ser.*, vol. 1751, no. 1, p. 012042, Jan. 2021, doi: 10.1088/1742-6596/1751/1/012042.

[27] C. Ding, T.-Y. Bao, and H.-L. Huang, “Quantum-inspired support vector machine,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 12, pp. 7210–7222, Dec. 2022, doi: 10.1109/TNNLS.2021.3084467.

[28] B. Zagajewski, M. Kluczek, E. Raczko, A. Njegovec, A. Dabija, and M. Kycko, “Comparison of random forest, support vector machines, and neural networks for post-disaster forest species mapping of the krkonoše/karkonosze transboundary biosphere reserve,” *Remote Sens.*, vol. 13, no. 13, p. 2581, Jul. 2021, doi: 10.3390/rs13132581.

[29] M. A. Bustomi and M. F. Asy’ari, “Analysis of cantaloupe fruit maturity based on fruit skin color using naive bayes classifier,” *J. Phys. Conf. Ser.*, vol. 1805, no. 1, p. 012028, Mar. 2021, doi: 10.1088/1742-6596/1805/1/012028.

[30] Y. A. Alsariera, V. E. Adeyemo, A. O. Balogun, and A. K. Alazzawi, “AI meta-learners and extra-trees algorithm for the detection of phishing websites,” *IEEE Access*, vol. 8, pp. 142532–142542, 2020, doi: 10.1109/ACCESS.2020.3013699.

[31] L. A. Demidova, “Two-stage hybrid data classifiers based on SVM and kNN algorithms,” *Symmetry (Basel).*, vol. 13, no. 4, p. 615, Apr. 2021, doi: 10.3390/sym13040615.

[32] A. Widodo and S. Handoyo, “The classification performance using logistic regression and support vector machine (SVM),” *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 19, pp. 5184–5193, 2017.

[33] M. Oltean, “Fruits 360,” 2020. [Online]. Available: https://www.kaggle.com/datasets/moltean/fruits

BIOGRAPHIES OF AUTHORS

Nashaat M. Hussien Hassan was born in Quena, Egypt, in 1977. He received his B.Sc. in communication and electronics engineering from Al-Azhar University-Egypt in 2002. In 2005, he received his M.Sc. degree in communication and electronics engineering from (C.N.M.) National Center of Microelectronics, Seville University-Spain. In 2009, he received his Ph.D. in Digital Integrated Circuit Design for the Applications of Image processing from (C.N.M.) National Center of Microelectronics, Seville University-Spain. In October 2019 he was promoted to the position of an Associate Professor position. Currently, he is working as an associate professor in the department of Electronics & Electrical communication, Faculty of Engineering, Fayoum University-Egypt. His research interest includes algorithms development (analysis, design, and improvement) and full-cycle software and Hardware product development (MATLAB, C, C++, VHDL, FPGA, and Xilinx). He can be contacted at email: nmh01@fayoum.edu.eg.

Basma Ramadan Gamal Elshoky was born in Minia, Egypt, in 1994. She received his B.Sc. in computer science from Minia University-Egypt in 2016. She accomplishes a pre-master in computer from Minia University in 2019. She is working as a teaching assistant in the Information technology section, Korean Egyptian Faculty for Industry and Energy Technology, Beni Suef Technological University-Egypt. She can be contacted at email: basma.r.gamal@gmail.com.

A. M. Mabrouk was born in Elmofya, Egypt, in 1988. He received his B.Sc. in electronics and electrical communications Menoufia University–Egypt in 2011. In 2019, he received his M.Sc. degree in electronics and electrical communications from Menoufia University-Egypt. In 2021, he received his Ph.D. in electronics and electrical communications from Minia University - Egypt. Currently, he is working as an assistant professor in the department of Electronics and Electrical communication, Faculty of Engineering and technology, Badr University in Cairo (BUC)-Egypt. His research interest includes electronics, communications, antenna, and microwaves. He can be contacted at eng_amosaad44@yahoo.com, ahmed.mosaad@buc.edu.eg.