International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

A Comparative Study of Multiclass Classification Using the Different Machine Learning Techniques for Fruit Species Prediction from Images

Rugved Korde[1], Achal Sultane[2], Vedant Mahitkar[3], Shreyal Ikhar[4], Achal Mokhale[5], Prof. S. N. Sawalkar[6]

[1,2,3,4,5] *U.G Students, Department of Computer Science and Engineering, SIPNA COET, Amravati, Maharashtra, India.*

[6]*Professor, Department of Computer Science and Engineering, SIPNA COET, Amravati, Maharashtra, India*

Abstract - Fruits play a vital role in our healthy life and are also used for the treatment of various diseases. It also contains an enormous quantity of fibers. It is the application of machine learning that we are using in the fruit classification model. Here we have different fruit images and we have to classify them using multiple algorithms. We are using various algorithms like KNN, random forest, Naive Bayes algorithms, etc. When we are using these algorithms we need our data in numbers or we can say in the numeric format, so we have to convert our fruit image data into a numeric format, and then by applying the various algorithms we can perform the task of classification. In this paper, a machine learning-based approach is presented for classifying and identifying different fruits with a dataset that contains various images. Some images are for training and some images are for validation and for testing. Here we have to take note of one thing while we are dealing with the machine learning and deep learning task or any project we want our data in numeric format. Here we are importing various types of libraries. Food security is a very important topic of discussion in today's society, as improper handling and management of food during production, processing, or distribution have caused increased food wastage around the globe. In addition, it has become clear from statistics gotten from surveys by institutions around the world like the Food Bureau of the United States that it is necessary to increase our rate of food production to meet the needs of our rapidly growing population.

***Key Words*:** Machine learning, image processing, prediction, classification, fruit, health, diseases

1.INTRODUCTION

In this paper, we are going to classify the fruit images by using machine learning algorithms. The fruits available naturally will be having different colors and shapes in appearance. Humans can identify the type of fruit by seeing their shape and color without any difficulty. Here a practical approach has been offered in this project to classify the fruit images based on the Color and Shape of the fruit.The images were taken from the standard dataset for the experiment; the dataset contains 28 types of fruit classes 'Grape Blue': 984, 'Pear': 696, 'Apple Braeburn': 492, 'Apple Granny Smith': 492, 'Apricot': 492, 'Cantaloupe': 492, 'Cherry': 492, 'Papaya': 492, 'Peach': 492, 'Pomegranate': 492, 'Strawberry': 492, 'Banana': 490, 'Cactus fruit': 490, 'Clementine': 490, 'Limes': 490, 'Mango': 490, 'Passion Fruit': 490, 'Pineapple': 490, 'Raspberry': 490, 'Orange': 479, 'Watermelon': 475, 'Kiwi': 466, 'Blueberry': 462, 'Corn': 450, 'Potato Red': 450, 'Plum': 447, 'Avocado': 427, 'Cucumber Ripe': 392. The color moment and shape of the fruits were considered to extract the features from different fruit images. In this proposed work, three feature vectors are constructed. In the color moment feature extraction, statistical features such as the mean and standard deviation of three-color channels (RGB) are computed here. The binarized images of fruits were used to extract shape-based features, and a multi-featured vector consisting of color moment and shape features was used.

In this project, We would be developing a simple training model to identify fruits in images. This would be a baseline effort for the development of a fruit classification system that can eventually be developed to identify bad fruits and vegetables and eventually be able to predict the multiclass of a fruit or vegetable using various classifiers. Machine learning is a powerful tool that has been applied to many fields for the purpose of automation of basic operations and optimization of the results of these operations. Research is ongoing into the application of machine learning for the inspection and grading of fruits and vegetables in retail stores to ensure accurate consistent reports on the quality of the products sold to consumers. In addition, seeing as the visual assessment is the primary basis for a purchase choice in

the market, ensuring the visual quality of fruits and vegetables is important to drive sales.

2. LITERATURE SURVEY

The way a fruit or vegetable looks is the first basis for which a purchase decision is made. The average person would not be able to track species and genetic traits of a fruit to be sure with any certainty whether a piece of fruit or vegetable is good or would taste sweet. To a good extent however, these conclusions drawn from the basic analysis of the fruit or vegetable by the customer can actually show the quality of the fruit which then means that the supplier or retailer needs to make sure that the fruits and vegetables they put on sale looks desirable, so as to drive up purchases. For this reason, A lot of time and effort is actually put into inspection and grading of fruits and vegetables at supermarkets and stores. For the sake of quality assurance and as a healthcare measure, certified inspectors also inspect and grade fruits and vegetables.

However, inconsistencies in their reports can also lead to high levels of wastage or bad purchases. Every false positive (that is bad fruits or vegetable graded as good) leads to a bad purchase, and every false negative (that is good fruit or vegetable graded as bad) leads to wastage. As a result of this, only about 60% of fruits and vegetables produced in the United States actually make it from the farm to the table (Farm Bureau, n.d.). Furthermore, fruit identification also reduces the time and effort needed for sorting of fruits at supermarkets and eliminates the need for direct contact with a lot of the farm produce along the supply chain. With new strains of viruses and bacteria causes health issues globally, being able to eliminate the unnecessary and inappropriate handling of farm produce by non-professionals through automation could help to solve problems of food poisoning among others.
A lot of fields and industries have seen the intervention of machine learning and taken advantage of its powerful automation and power to increase efficiency and reduce cost and effort needed. As a result, these areas have had their basic operations revolutionized. Although there have been a few areas that have seen more research and work done by artificial intelligence, there are a few other areas that are also beginning to pick up. One of these areas is Agriculture.

Fruit and vegetable recognition by fusing color and texture features was proposed by [7]. Their design relies on the background subtraction and classification with a multi-class support vector machine. [17] used fusion of multiple color channels for multi-class fruit detection from pictures with highly complex backgrounds. However, the bottleneck is the detection speed which takes 28 seconds to process one image. Color and texture features were derived for fruit recognition using the minimum distance classifier [2]. The texture features are based on the sum and difference histogram of the neighboring pixels and are presented by [8]. With an exception of [17], all the other papers used images of fruits and vegetables on a white background. They report that the main discriminative features between the different classes of fruit and vegetable are based on color and texture.

3. METHODOLOGY

a) Home dashboard

A simple interface after opening the application. which will instruct the user to select the fruit image form the device storage. In this way application frontend forward the image data towards backend.

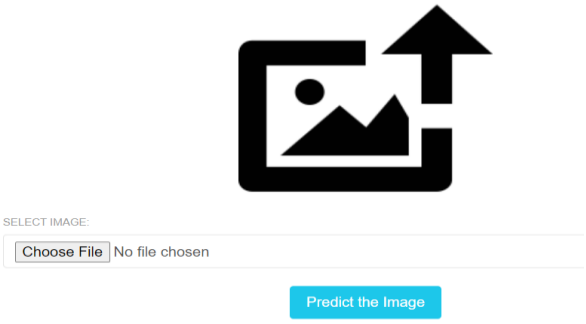

Fig1: Home Dashboard

b) Fruit prediction

After browsing the image there is a button named predict the image which will display the final output given by the backend flow.

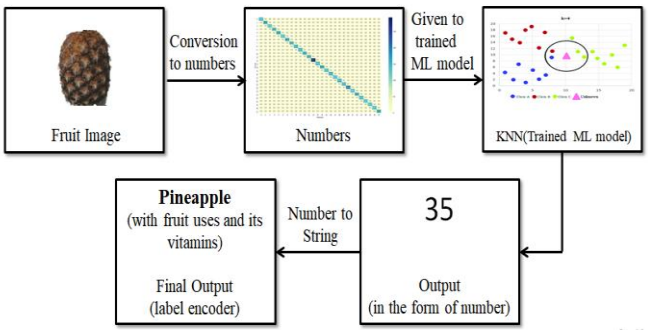

Fig2: Application's Backend Flow

c) KNN(K-Nearest Neighbour)

KNN refers to K-Nearest Neighbour algorithm, KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data. K-NN algorithm assumes the similarity

between the new data and available cases and put the new case into the category that is most similar to the available categories. It is the non parametric learning algorithm.

KNN works by finding the distances between a query and all the examples in the data, selecting the specified number examples (K) closest to the query, then votes for the most frequent label

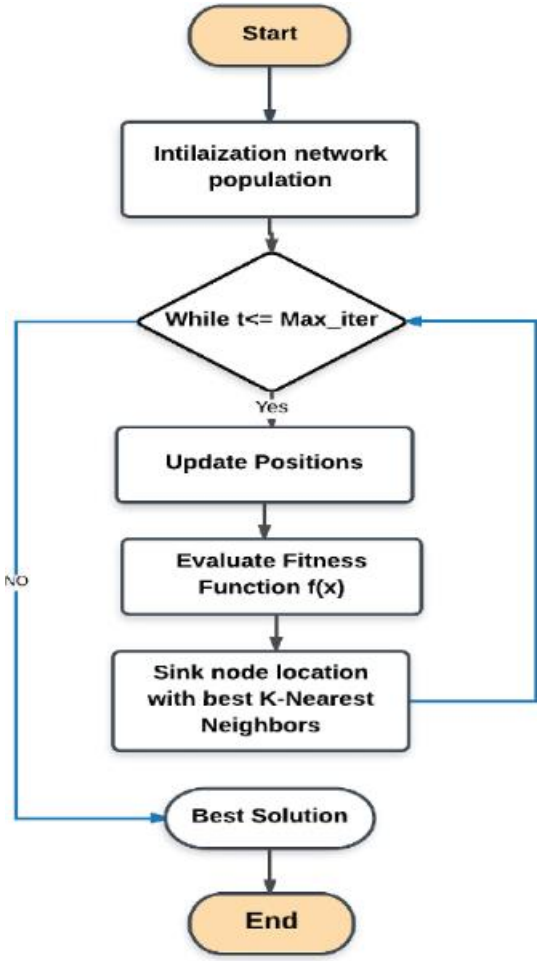

Fig3: KNN Algorithm workflow

4. SYSTEM DESIGN

System design is the most important and vital part of any framework as it is used for the development of the system from its theory. This section includes the modules, architecture and various elements that are combined together to form the whole systems framework.

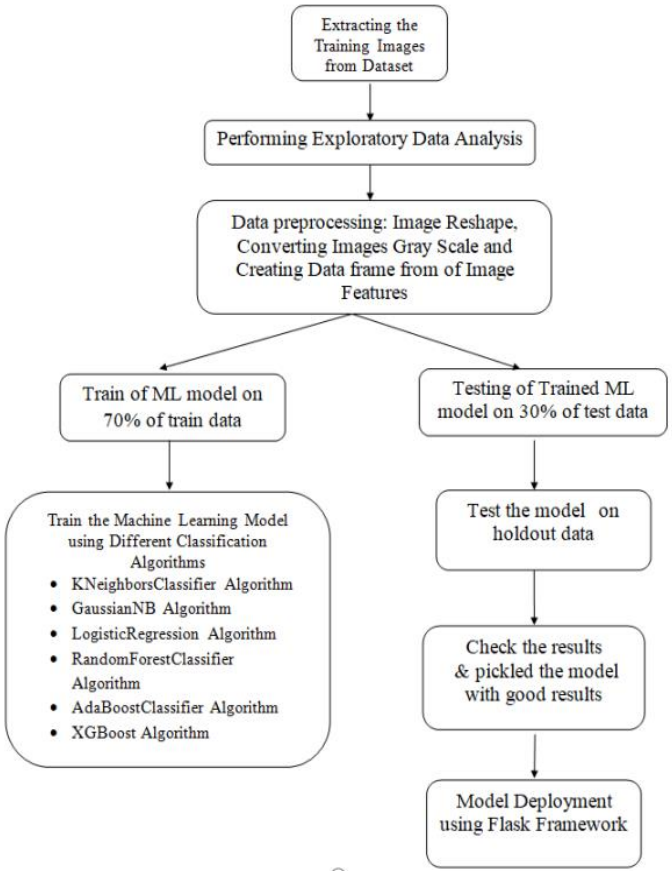

Fig4: Application Workflow

This proposed system is based on Machine Learning and this system offers a useful method for categorizing fruit photos based on the colour and shape of the fruit. In this study, a classification unit and image processing algorithm were created to categories the fruits based on their size and colour features. for this reason. To ascertain the success rates of categorising the fruits, the calliper and spectrophotometer measurements of size and colour were input into the proposed image processing system. the combination of the classification unit and the image processing method On the other hand, data mining techniques were employed to analyse the size and colour values that the image processing system read from the fruits. Fruit classification remains a challenge because of the similarities involved by large quantities of types of fruits. With the aim of recognizing fruits accurately and efficiently, this research offered a novel fruit classification model.

The proposed methodology consisted of the following four processes: (i) A four-step preprocessing was performed. (ii) The color, shape, and texture features were combined. (iii) Principal component analysis was employed for feature reduction. (iv) We presented various classification methods. In this way we obtain our final results.

It helps the patients to recognize the fruit and its uses like vitamins and foe which disease the one can consume such fruit.

5. RESULTS AND DISCUSSION

The final results of the application shows the accuracy score based on the analysis of different machine learning algorithms used. And the combined feature vector with the KNN classifier has received 99.98% recognition accuracy so we applied KNN classifier to this system. Following is the accuracy table of the different classifiers discussed:

Algorithm	KNN	GaussianNB	Logistic Regression	Random Forest	Adaboost	XGboost
Acc of Classifier	99.952			89.936		
Accuracy	0.9995	0.91096	0.99857	0.8993	0.17452	0.9992
Error Rate	0.000473	0.8903	0.001420	0.10063	0.82547	0.000710
Precision	0.99952	0.92621	0.99861	0.91864	0.13609	0.9992
Recall	0.999526	0.91096	0.99857	0.89936	0.17452	0.99928
F1 Score	0.99952	0.9123	0.99858	0.88505	0.13296	0.99928

Fig5: Tabular representation of accuracy rate of classifiers

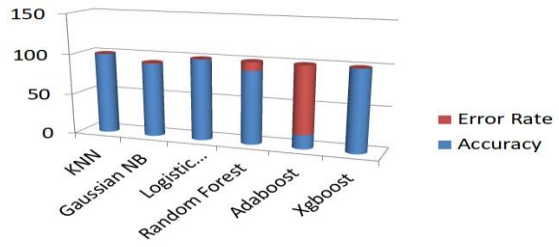

Fig6: Graphical representation of accuracy rate of classifiers

6. CONCLUSIONS AND FUTURE SCOPE

CONCLUSIONS

Sometimes the one feature descriptor is not sufficient to capture the classes' dissimilarity when it comes to solving complex classification issues. Efficient and successful fusion policies are, therefore required, and it becomes essential. Hence, this project successfully utilized the two feature vectors with the fused feature vector. The multiple classifiers have been considered to obtain the highest recognition accuracy. The combined feature vector with the KNN classifier has received 99.98% recognition accuracy. More features may be considered in future work, and an attempt has been to apply deep learning techniques on the fruit images.

FUTURE SCOPE

A few organizations like Amazon and some large-scale farms have been able to see the value of including machine learning to aid in their daily operations. This application has shown many benefits to them and helped them reduce cost and increase efficiency, however, there is still more that can be done.
Computer vision experts at Amazon are still doing research into how this system can be improved to predict when a fruit is going to go bad and to tell if a fruit would be sweet. This would require much more data and computation power, the like that provided by Amazon with the Amazon Web Services (AWS).
In conclusion, Fruit classification is an important research are that should be looked into and have more research and development resources because its development and improvement can have a far-reaching effect on agriculture and the quality of fruits provided at markets and retail stores.

7. ACKNOWLEDGEMENT

It is my utmost duty and desire to express acknowledgement to the various torchbearers, who have rendered valuable guidance during the preparation of my Project. First, I extend my deepest gratitude to respected Principal, **Dr. S. M. Kherde** without whose support, my Project could not have been transformed into present form.

I am grateful to **Dr. V. K. Shandilya** Head, Computer Science and Engineering Department, and my guide **Prof. S. N. Sawalkar** for providing immense support and guidance. I am beholden for guiding me at every step in the Project. He has honestly guided me throughout, never leaving me unanswered for any of my doubts. It was his constant persuasion, encouragement, inspiration and able guidance that helped me in completing my Project successfully.

8. REFERENCES

[1] A. Nosseir and S. E. Ashraf Ahmed, "Automatic classification for fruits' types and identification of rotten ones using k-NN and SVM," Int. J. online Biomed. Eng., vol. 15, no. 3, pp. 47–61, 2019, doi: 10.3991/ijoe.v15i03.9832.

[2] S. K. Behera, A. K. Rath, A. Mahapatra, and P. K. Sethy, "Identification, classification & grading of fruits using machine learning & computer intelligence: a review," J. Ambient Intell. Humaniz. Comput., no. Kondo 2010, 2020, doi: 10.1007/s12652-020-01865-8.

[3] C. C. Patel and V. K. Chaudhari, Comparative Analysis of Fruit Categorization Using Different Classifiers. Springer Singapore.

[4] S. Sakib and Z. Ashrafi, "Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers," pp. 8–11, 1980.

[5] H. M. Zawbaa, M. Hazman, M. Abbass, and A. E. Hassanien, "Automatic fruit classification using random forest algorithm," pp. 164–168, 2014.
Journal of University of Shanghai for Science and Technology ISSN: 1007-6735 Volume 22, Issue 12, December - 2020 Page-1354

[5] A. Rocha, D. C. Hauagge, J. Wainer, and S. Goldenstein, "Automatic fruit and vegetable classification from images," vol. 70, pp. 96–104, 2010, doi: 10.1016/j.compag.2009.09.002.

[6] S. P. Deenan and J. Satheeshkumar, "A Study on Image Processing Methods for Fruit Classification A Study on Image Processing Methods for Fruit," no. December 2012, 2016.

[7] "RECOGNITION ALGORITHMS FOR DETECTION OF APPLE FRUIT IN AN ORCHARD FOR EARLY YIELD PREDICTION Rong Zhou ."

[8] C. Pl, DEFECT IDENTIFICATION IN THE FRUIT APPLE USING K-MEANS COLOR IMAGE SEGMENTATION ALGORITHM," no. August 2017, 2019, doi: 10.26483/ijarcs.v8i8.4735.

[10] J. P. Mercol, J. Gambini, and J. M. Santos, "Automatic classification of oranges using image processing and data mining techniques," XIV Congr. Argentino Ciencias la Comput. XIV Argentine Congr. Comput. Sci. (CACIC 2008), pp. 1–12, 2008.