International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Apple Fruit Disease Classification using Naive Bayes Algorithm

Naveen S, Sabarigireeson A T

Naveen-Artificial Intelligence and Machine Learning ,-Sri Shakthi Institute of Engineering and Technology
Sabarigireeson- Artificial Intelligence and Machine Learning ,Sri Shakthi Institute of Engineering and Technology

---***---

Abstract - Apples are one of the most productive varieties of fruit in the world, with a high nutritional and medicinal value. However, numerous diseases affect apple production on a wide scale, resulting in significant economic losses. These diseases often go overlooked until just before, after, or after fruit has been processed. Many pathogens can be avoided with cultural traditions and (optional) fungicides, even if there are no cures for tainted fruit. However, accurate diagnosis is essential for determining the right management practices and preventing further losses. Apple scab, apple rot, and apple blotch are some of the most prevalent diseases that affect apples. The proposed approach will greatly aid in the automated identification and classification of apple diseases, according to our test results. We discovered that normal apples were easy to discern from diseased apples in our trial, and that the texture-based GLCM function produced more reliable results for apple disease classification, with a classification accuracy of more than 96.43 percent. This demonstrates that combining the GLCM extraction function with naive bayes classification will greatly improve accuracy.

***Key Words*:** Naive bayes, Apple Diseases, Classifier

1.INTRODUCTION

Apple is a fruit that is eaten and cultivated all over the world, Because of its delicacy and high nutritional value , Apples are one of the most productive varieties of fruit in the world, with a high nutritional and medicinal value. However, numerous diseases affect apple production on a wide scale, resulting in significant economic losses . Traditionally, plant disease severity has been determined by qualified experts visually inspecting plant tissues, This results in high costs and inefficiency. Cultivation and management specialist systems have been widely used as a result of the widespread use of digital cameras and the advancement of information technology in agriculture, significantly increasing plant production capacity. Apple fruit diseases can result in major yield and quality losses. These diseases often go overlooked until just before, after, or after fruit has been processed. Many pathogens can be avoided with cultural traditions and (optional) fungicides, even if there are no cures for tainted fruit. However, accurate diagnosis is essential for determining the right management practices and preventing further losses. Healthy recognition of fruits and apples is an important issue for the economic and agricultural fields. The traditional method for detecting and identifying fruit diseases relies on professional examination through the naked eye. Owing to the remote areas of their supply, consulting experts can be costly and time consuming in some developed countries.Automatic identification of fruit diseases is critical in order to diagnose disease symptoms as soon as they occur on growing fruits. Fruit diseases can result in significant yield and quality losses when they occur during harvesting. Apple scab, apple rot, and apple blotch are some of the most prevalent diseases that affect apples. Apple scabs appear as gray or brown corky patches on the apple. Apple rot causes somewhat sunken, oval brown or black patches that are often surrounded by a red halo. Apple blotch is a fungus that causes black, uneven, or lobed edges on the fruit's surface. The following are three common apple fruit diseases: (a) apple scab, (b) apple rot, and (c) apple blotch Apple samples with identical feature values may be grouped into one group in a multidimensional space . The identification of apple diseases using the Plant Pathology Apple Dataset is discussed in this article. Pre- processing is the next stage, which includes image enhancement and grayscaling. The pre-processed image is then used to apply feature extraction methods. The Gray Level Co-occurrence Matrix (GLCM) is used in this case, with six parameters (Angular Second Moment, Contrast, Entropy, Variance, Correlation, and Inverse Different Moment). After that, the image generated by GLCM's feature extraction will be processed using two different learning algorithms: Naive Bayes

2. LITERATURE REVIEW

In previous research has been conducted a lot of research in the field of apple disease, but still not get significant results, here are some studies related to the problems of this study. Jamdar and Patil (2017) Presented Image processing based solutions are proposed and evaluated for the detection and classification of apple fruit diseases. The proposed approach consists of mainly three steps. In the first step, image segmentation is performed using the K-Means grouping technique. In the second step the features are extracted. In the third step, the training and classification was carried out at LVQNN and resulted in an accuracy of up to 90%. Misigo and Miriti (2016) Presented investigated the application and performance of Naive Bayes algorithm in the The 9 th International Conference on Cyber and IT Service Management (CITSM 2021) Bengkulu, September 22-23, 2021 classification of apple varieties. Comparison of their classification accuracy results with Naive Bayes technique shows that Naive Bayes accuracy is higher than the accuracy of analysis of major components, fuzzy logic and MLP- Neural with 91%, 90%, 89%, and 83% respectively. This research shows that Naive Bayes has good potential for non- destructive and accurate identification of apple varieties. Mishra and Barskar (2021) Presented In the future, classification techniques in machine learning such as decision trees, Naïve Bayes classifiers can be used for disease detection in plants and in the sense of helping farmers automatic detection of all kinds of diseases in plants to be detected because it has a very high accuracy . Yuan (2018) Presented Analysis of the Naive Bayes algorithm for image classification using the co-occurrence matrix texture attribute extraction approach yielded a 96 percent accuracy score. Yuan (2020) Presented This is the driving force behind the author's decision to use the algorithm to diagnose apple disease. Apple Disease Detection Using a Deep Learning Algorithm had an accuracy of 88.5 percent and For the automated identification and classification of fruit diseases, the CLBP function and Multi-class Support Vector Machine as a classifier were used. 93 percent of the time in previous studies. For the classification of apple disease, we used the GLCM function extraction and the naive Bayes algorithm in our research. It is anticipated that using GLCM and Naive Bayes would result in greater precision.

3.METHOD

Precision image segmentation is expected for the fruit disease classification problem; otherwise, the features of the non-infected region would dominate over the features of the infected region. K-Means dependent image segmentation is favored in this method to detect the area of interest, which is only the infected portion. Features are derived from the segmented image of the fruit after it has been segmented using GLCM. Finally, a naive bayes classifier is used for training and classification. The suggested solution's structure is depicted. The rest of this section goes over each phase of the proposed method. Framework of the proposed approach.

A. Data acquisition:

Data acquisition was done to obtain the sampling data needed in this study. The object used is using 4 types of apple disease,namely apple blotch, apple rot, apple scab and healthy.

B. Image Segmentation:

At this stage to get an image with the same provisions with the intention of making it easier for the system to process the image is done cropping and then the image data used is converted into grayscale imagery with the aim to be able to display 1 color on each image after that is done noise removal to smooth the image so that in the process later can produce maximum results that are then done detection of canny edge lines that are useful as segmentation c itra. on each preprocess done with the intention to give good results at the time of extracting texture characteristics.

C. Feature Extraction:

Grey Level Co-occurrence Matrix Method is a matrix whose components are the number of pairs of pixels with the same brightness degree, divided by d pixels and with an inclination angle of. In other words, the matrix is the likelihood of the two pixels separated by d and angle j having the same gray level I and j. As seen in the figure, neighboring pixels with a distance of d between them can be found in eight different directions. Haralick et al. suggest different types of statistical texture properties that can be derived using the co-occurrence matrix approach. Any of these include, but are not limited. The four characteristics of GLCM are extracted and summarized as follows: comparison, entropy, homogeneity, and energy. The following is the equation for these characteristics:

1. In the GLCM matrix, the gray level variation is visible as contrast. It calculates the pixel's and its neighbor's intensity.
2. The energy function is used to calculate local homogeneity, which is referred to as the Entropy. It has a spectrum of 0 to 1 as a value.
3. Homogeneity feature: in the GLCM, compute the not-zero, which is the opposite of the contrast weight. It has a spectrum of 0 to 1 as a value.
4. The amount of energy, or entropy.

D. Naïve Bayes

Classification is a work training/learning to target function f that maps each set of attributes (features) x to one of the available class labels y. Job teaching would result in the creation of a blueprint, which will then be stored as a memory[. To construct a model, classification algorithms use training data. The model is then used to estimate the class mark of new data that has not yet been classified. Democracy (independence) is a solid (naive) assumption. The "model with individual functions" is the model that was used. The application of Bayes' theorem is used to create a basic probabilistic-based prediction strategy known as Naive Bayes.

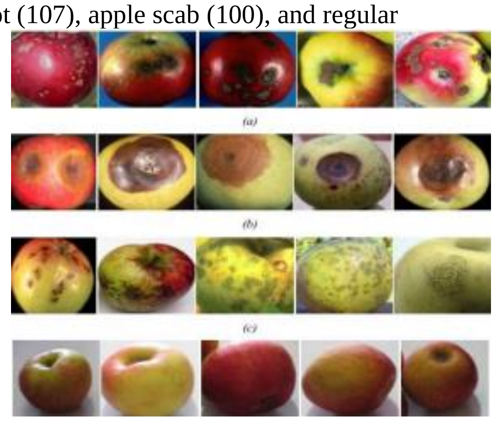
Democracy (independence) is a solid (naive) assumption. The 9 th International Conference on Cyber and IT Service Management (CITSM 2021) Bengkulu, September 22-23, 2021

The "model of individual features" is the model that was used. The easiest approach for using current opportunities is the Nave Bayes classification, which assumes that any vector X is open (independence)

4. RESULT AND DISCUSSIONS

A.Datasets

A dataset of diseased apple fruits was also examined, with 391 photographs divided into four categories: apple blotch (104), apple rot (107), apple scab (100), and regular



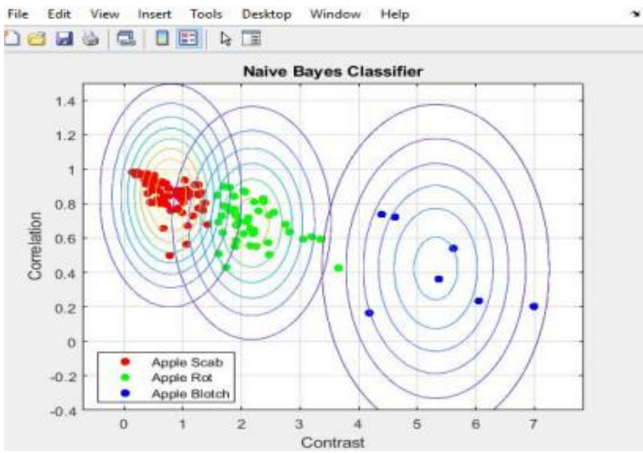
B. Feature Extraction

We get GLCM values at each angle for an apple picture from GLCM in Algorithm. To simplify the interpretation of the proposed procedure, we took one of the closest samples to display the findings clearly, the three disease pictures. When we add the apple image to the Gray level co-occurrence matrix in our proposed process, we get eight values for each angle (0o, 45o, 90o, and 135o), which we use in the texture feature extraction stage to get four features. Figure 5 shows the GLCM result for an apple image, which will be used in the feature

	0	45	90	135	average
Contrast	0.039378	0.055427	0.036399	0.057972	0.047294
Correlation	0.99268	0.98969	0.99322	0.98922	0.9912
Energy	0.19458	0.19027	0.19547	0.18987	0.19255
Homogeneity	0.9807	0.97307	0.98226	0.97236	0.9771

C. Classifier

In this study using matlab as supporting software for the implementation of naive bayes the following is a display of naive bayes algorithm modeling using matlab



The picture above is a modeling of the Naive Bayes algorithm for disease classification using Matlab, first the dataset to be processed is entered into the Citra_Latih folder, then the dataset is processed using the GLCM feature feature and then for classification using the Naive Bayes algorithm to get clusters of 3 diseases from apples.

5. CONCLUSION

In this article, image processing-based solutions for detecting and classifying apple fruit disease are suggested and evaluated. The suggested method consists primarily of four stages. Data preprocessing is performed in the first step, followed by image segmentation in the second. GLCM is used to remove the functionality in stage three. The fourth stage included using Naive Bayes for preparation and classification. As case studies and to test our software, we used three forms of apple disease: Apple Blotch, Apple Rot, and Apple Scab. The proposed approach will greatly aid in the automated identification and classification of apple diseases, according to our test results. We discovered that normal apples were easy to discern from diseased apples in our trial, and that the texture-based GLCM function produced more reliable results for apple disease classification, with a classification accuracy of more than 96.43 percent. This demonstrates that combining the function with naive bayes classification will greatly improve accuracy.

6.ACKNOWLEDGEMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page

7.REFERENCES

[1] J. P. An, X. W. Zhang, R. R. Xu, C. X. You, X. F. Wang, and Y. J. Hao, "Apple MdERF4 negatively regulates salt tolerance by inhibiting MdERF3 transcription," Plant Sci., vol. 276, pp. 181–188, 2018, doi: 10.1016/j.plantsci.2018.08.017. [2] D. Zhang et al., "Nondestructive measurement of soluble solids content in apple using near infrared hyperspectral imaging coupled with wavelength selection algorithm," Infrared Phys. Technol., vol. 98, no. December 2018, pp. 297–304, 2019, doi: 10.1016/j.infrared.2019.03.026. [3] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks," IEEE Access, vol. 7, pp. 59069–59080, 2019, doi: 10.1109/ACCESS.2019.2914929. [4] M. Dutot, L. M. Nelson, and R. C. Tyson, "Predicting the spread of postharvest disease in stored fruit, with application to apples," Postharvest Biol. Technol., vol. 85, pp. 45–56, 2013, doi: 10.1016/j.postharvbio.2013.04.003. [5] B. Liu, Y. Zhang, D. J. He, and Y. Li, "Identification of apple leaf diseases based on deep convolutional neural networks," Symmetry (Basel)., vol. 10, no. 1, 2018, doi: 10.3390/sym10010011. [6] N. W. Gauthier, "Fruit Diseases of Apple," Coll. Agric. Food Environ., 2019, [Online]. Available: https://plantpathology.ca.uky.edu/files/ppfs-fr-t- 02.pdf. [7] S. A. Gholve and P. M. A. P, "A Fruit Identification with Classification Fault Detection Technique using K-means clustering," Int. J. Sci. Dev. Res., vol. 5, no. 11, pp. 115–120, 2020. [8] S. R. Dubey and A. S. Jalal, "Detection and classification of apple fruit diseases using complete local binary patterns," Proc. 2012 3rd Int. Conf. Comput. Commun. Technol. ICCCT 2012, no. December, pp. 346–351, 2012, doi: 10.1109/ICCCT.2012.76. [9] Y. Yu, S. A. Velastin, and F. Yin, "Automatic grading of apples based on multi-features and weighted K-means clustering algorithm," Inf. Process. Agric., vol. 7, no. 4, pp. 556–565, 2020, doi: 10.1016/j.inpa.2019.11.003. [10] S. Saifullah and V. A. Permadi, "Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms," Proceeding - 2019 5th Int. Conf. Sci. Inf. Technol. Embrac. Ind. 4.0 Towar. Innov. Cyber Phys. Syst. ICSITech 2019, pp. 140– 145, 2019, doi: 10.1109/ICSITech46713.2019.8987496. [11] E. Wijaya, "Implementation Analysis of GLCM and Naive Bayes Methods in Conducting Extractions on Dental Image," IOP Conf. Ser. Mater. Sci. Eng., vol. 407, no. 1, 2018, doi: 10.1088/1757- 899X/407/1/012146. [12] G. Xuan et al., "Apple Detection in Natural Environment Using Deep Learning Algorithms," IEEE Access, vol. 8, pp. 216772–216780, 2020, doi: 10.1109/ACCESS.2020.3040423. [13] A. V Jamdar, "Detection and Classification of Apple Fruit Diseases using K-means clustering and Learning Vector Quantization Neural Network," vol. 2, no. 6, pp. 423–429, 2017. [14] R. Misigo, "Classification of Selected Apple Fruit Varieties Using Naive Bayes," Univ. Nairobi, vol. 3, no. 1, p. 56, 2016, [Online]. Available: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/Graue Publikationen/MT_Globalization_Report_2018.pdf %0Ahttp://eprints.lse.ac.uk/43447/1/India_globalisa tion%2C society and inequalities%28lsero%29.pdf%0Ahttps://www.quora.com/What-is-the. [15] N. Prashar, "A Review on Plant Disease Detection Techniques," no. March, pp. 501–506, 2021, doi: 10.1109/icsccc51823.2021.9478132. [16] S. A. Alazawi, N. M. Shati, and A. H. Abbas, "Texture features extraction based on GLCM for face retrieval system," Period. Eng. Nat. Sci., vol. 7, no. 3, pp. 1459–1467, 2019, doi: 10.21533/pen.v7i3.787. [17] Prasetyo E, Data Mining Konsep dan Aplikasi Menggunakan Matlab. Yogyakarta: Andi, 2002. [18] S. R. Dubey and A. S. Jalal, "Adapted approach for fruit disease identification using images," Image Process. Concepts, Methodol. Tools, Appl., vol. 3–3, pp. 1395–1409, 2013, doi: 10.4018/978-1-4666- 3994-2.ch069.

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1 (1997) 108–121
2. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): Theoretical Aspects of Computer Software. Lecture Notes in Computer Science, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York (1997) 415–438
3. van Leeuwen, J. (ed.): Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)
4. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)