# Detection And Classification of Various Diseases of Apples In Leaves And Fruits Using Machine Learning: A Review

Shafiya Mushtaq, Dr. Ashish Oberoi
M.Tech Student, Department of Computer Science & Engineering, RIMT University
Professor, Department of Computer Science & Engineering, RIMT University
*(shafiya9.sm@gmail.com)*

***Abstract—*** The presence of diseases in various kinds of fruits is the major factor of production and the economic degradation of the agricultural industry worldwide. In current scenarios there are too many problems for agriculture sector like pollution, shortage of water, global warming etc. Because of them there are plenty of diseases which are spoiling the crops grown by farmers. Here, our idea is to provide a smart solution to save crops from such diseases. An approach for various apple disease classification and detection using color, texture and shape-based attributes is explored and experimentally verified in this paper. The basic steps of the introduced image processing-based method are (a) disease detection is done with the help of K-means clustering method,(b)color, texture and shape-based features are determined over the segmented image and combined to form the single descriptor, and (c) multi-class support vector machine is used to classify the apples into one of the infected or healthy categories. Apple fruit is taken as the test case in this study with many categories of diseases, namely, marsonina leaf blotch, apple rot and apple scab,cedar-apple rust as well as healthy apples. A review of all different classifications and detection techniques was done and hence it was concluded that K means clustering is best.

***Keywords—*** *Disease detection, Image gray scaling, K-Means Clustering, Support Vector Machine, Color, Texture, Shape, Local Binary Pattern.*

## I. INTRODUCTION

Diseases in fruits cause devastating problem in economic losses and production in agricultural industry especially in country like India where most of the revenue is generated from agriculture. Apple production in India is limited to few states such as Jammu & Kashmir, Himachal Pradesh and Uttarakhand. And these are the states where people cannot get instant help if crop has any kind of disease because of the various reasons like bad weather, lack of fast transportation facility etc. which effects more to the producer in case of diseased fruits. This situation is further complicated by the fact that today diseases are transferred globally more easily than ever before. [1] New diseases can occur in places where they were previously unidentified where there is no local expertise to combat them. Fewer Knowledge of pesticide usage can cause the development of long-term resistance of the pathogens, reducing the ability to fight back.

Timely and accurate diagnosis of plant diseases is most important for precision agriculture. Because of these diseases considerable amount of fruit is destroyed. So diseases of apples are major concerns for its farmers.

Apple commonly has three diseases apple scab, apple rot, and apple blotch as shown in Fig 1. These diseases can be identified by each unique symptoms. Apple scabs are grayish or brown corky spots. Apple scab spreads rapidly between 50-75 degrees. Apple rot is slightly sunken, circular brown or black spots which may be covered by a red halo. Apple blotch diseases are fungal diseases that appear on the surface of the apple fruit as dark, irregular or lobed edges. Therefore, we should have some sort of computerized system which could identify these diseases by the symptoms visible in the leaf and fruit image. This way we can easily and early determine the disease and we can get solutions too. Symptoms on leaves are similar to those found on fruits. Hence, disease prediction in apple cultivation is of considerable importance to overcome these problems. Some diseases also infect other parts of the tree causing diseases of twigs, leaves, and branches.[2]

The classical approach for identification of various diseases is based on the naked eye observation by the experts. In some developing countries, consulting experts are very expensive and time consuming due to the distant locations of their availability. Automatic detection of diseases is very important to automatically detect the symptoms of diseases as early as they appear on the growing fruits and leaves. Apple fruit diseases can cause major losses in yield appeared in harvesting. To know what control factors to take next year to avoid losses, it is crucial to recognize various factors that cause diseases. In this paper, we propose and experimentally assign an adaptive approach for the identification of fruit diseases using images.

With the advent of technology, we are trying to make a system that can think like humans, so we explored a new field called artificial intelligence. Machine learning constitutes a recent, modern technique for image processing and data analysis, with accurate results and large potential [3]. As machine learning has been successfully applied in various domains, it has also recently entered in the domain of agriculture. So we will apply machine learning to create an algorithm for automated detection and classification of various crop diseases. The

proposed approach is composed of the following steps; in first step the images are segmented using K-Means clustering technique, in second step, some features are extracted from the segmented image, and finally, diseases are classified using a Multi-class Support Vector Machine (SVM). The early detection of plant diseases could be a valuable source of information for executing proper diseases detection, plant growth management strategies and disease control measures to prevent the development and further spread of diseases. [4]

Fig 1. (Various diseases of apples)

## II. RELATED WORK

Konstantinos et.al. (2018) used CNN to classify the plant leaf disease using both diseased and healthy plant. They have tested with AlexNet, AlexNetOwtBn, GoogleNet, Overfeat and VGG using Torch7. They compared original image and pre-processed image for 5 different CNN architectures. They found that the accuracy rate is 99.44%. They have also proved that CNN is very much suitable for the early stage detection of plant diseases. [4]

BinLiu et.al (2017) used a methodology to identify some apple leaf diseases. They generated pathological images of apple leaf disease and found the results for four types of leaf diseases. They had evaluated with 13,689 leaf images and identified the disease at the early stage in order to avoid the spread of infection. They achieved 97% of accuracy recognition rate and this result is increased by 10% of accuracy compared to the convolutional models. In future we can implement FasterRCNN (Regions with CNN), and SSD (Single Shot MultiBox Detector) algorithms. [5]

Manuel Cortes et.al. (2017) used 84,147 images of diseased and healthy plants to classify crop species and disease status of 57 different classes using deep convolutional network and semi supervised methods. They have used GLCM functions, Lacunarity and shen features to characterize texture of image by calculating specific values and spatial relationship that occur in an image for pairs of pixel. Using HSV features and SVM, they have also proposed a novel neural network architecture to classify the infected leaf and healthy leaf. After classification they have implemented genetic algorithm to optimize SVM loss and identify the disease type in the infected leaves. Their combination of this algorithm never works in this research field. However, for unstructured data the performance is imported to 12%. [6]

Srdjan Sladojevic et.al. (2016) used a novel methodology to facilitate a quick and easy implementation of plant disease recognition to recognize 13 different types of plant diseases. They have performed Deep CNN training using a Deep Learning framework named Caffe and achieved an accuracy of 96.3%. After fine tuning of parameters in 100th training. The images taken from the dataset has undergone 4 layer filtering technique using CNN. The proposed model can automatically classify and detect 13 different plant diseases using 3,000 original images later extended up to 30,000 images using appropriate transformation. [7]

Vipinadas.M.J et.al. (2016) have proposed a novel method for image pattern classification in banana leaf using SVM and ANFIS classifier. They have graded banana leaves using ANFIS and compared the performance of SVM and ANFIS classifiers. Using multiple levels SVM they have classified the diseases as blacksigatoka and panamawait in banana leaves. In this classification process first RGB color image is converted into Ycbcr color space and grey scale image. [8]

Srdjan Sladojevic et.al. (2016) have proposed a novel methodology to facilitate a quick and easy implementation of plant disease recognition to recognize 13 different types of plant diseases. They have performed Deep CNN training using a Deep Learning framework named Caffe and achieved an accuracy of 96.3%. After fine tuning of parameters in 100th training. The images taken from the dataset has undergone 4-layer filtering technique using CNN. The proposed model can automatically classify and detect 13 different plant diseases using 3,000 original images later extended up to 30,000 images using appropriate transformation. [9]

Miss.Kambale, Anuradha Manik, Dr.Mrs.Chougule et.al (2015) they used Image processing for classification and grading of apple fruit images. According to this paper, first step is to create two databases one with normal images and one with defected images. Second step is image pre-processing and image segmentation. Third step is feature extraction and after that classification and grading. experimental result of this algorithm indicate that the proposed solution and can be significantly support automatic classification and grading of apple fruit diseases. [10]

Shivaram Dubey, Anand Singh Jalal et.al (2012) used three apple diseases, apple rot and apple blotch. In detection of the apple disease by image the first step is image segmentation by K-mean clustering. In second step features are extracted. These features are extracted from the segmented image and that features are global color histogram (GCH), color

coherence vector (CCV), local binary pattern (LBP) and complete local binary pattern (CLBP). In which complete local binary pattern give 93% correct classification accuracy than other features because it calculates magnitude, sign and center value of pixel. [11]

Lopez-Garcia et al. (2010) developed a citrus fruits surface detection algorithm, which was based on multivariate image analysis strategy and principal component analysis approach. The success ratio for the detection of the individual defect was 91.5%. The classification accuracy was acceptable. However, the complexity of the algorithm restrained the detection speed, which made it impractical to be used online. [12]

Kim et al. (2009) used 14 colors, texture features of grapefruit peel diseases to classify defects based on HSI (Hue, Saturation, and Intensity) model [6]. Cai et al. (2009) used hyperspectral imaging technology to detect rust in citrus in this study, and band ratio algorithm was proposed to overcome the adverse effects of uneven reflectance intensity due to the curvature of spherical objects. [13]

Blasco et al. (2007) developed an unsupervised growing region algorithm to detect citrus surface defects; the algorithm finds two regions in a fruit with peel defects: the larger region corresponded to the sound peel and smaller areas were defects. [14]

Marcassa et al. (2006) have applied laser induced fluorescence spectroscopy to investigate biological processes in orange trees. They have investigated water stress and Citrus Canker, which is a disease produced by the Xanthomonas axonopodis pv. citri bacteria. They have discriminated the Citrus Canker`s contaminated leaves from the healthy leaves using a more complex analysis of the fluorescence spectra. However, they were unable to discriminate it from another disease. [15]

Bravo et al. (2004) used fluorescence imaging multispectral or hyperspectral imaging used by Moshou et al. (2006), infrared spectroscopy used by Spinelli, Noferini, & Costa (2006), visible/multiband spectroscopy used by Yang, Cheng, & Chen (2007), Chen et al. (2008), and nuclear magnetic resonance (NMR) spectroscopy used by Choi et al. (2004). [16]

Li, Wang, & Gu; Mehl et al. 2002 performed defect segmentation of fruits are done using simple threshold approach A globally adaptive threshold method (modified version of Otsu`s approach) to segment fecal contamination defects on apples are presented by Kim et al. (2005). Classification-based methods attempt to partition pixels into different classes using different classification methods.Bayesian classification is the most used method by researchers Kleynen, Leemans, & Destain (2005) and Leemans, Magein, & Destain (1999), where pixels are compared with a pre-calculated model and classified as defected or healthy. Unsupervised classification does not benefit any guidance in the learning process due to lack of target values. [17]

## III. RESEARCH METHODOLOGY:

K -means clustering algorithm is an unsupervised algorithm technique and it is used to segment the interested area from the background.K-means clustering technique is used for the defect segmentation. Images are grouped into four clusters in which one or more cluster contains only infected region of the fruit as shown in Fig 2. K-means clustering algorithm was developed by J. Mac Queen (1967) and later by J.A.Hartigan & M. A. Wong . The K-means clustering algorithms classify the objects (pixels ) into K-number of classes based on a set of features. The classification is done by minimizing the sum of squares of distances between the data objects and the corresponding cluster. [18]

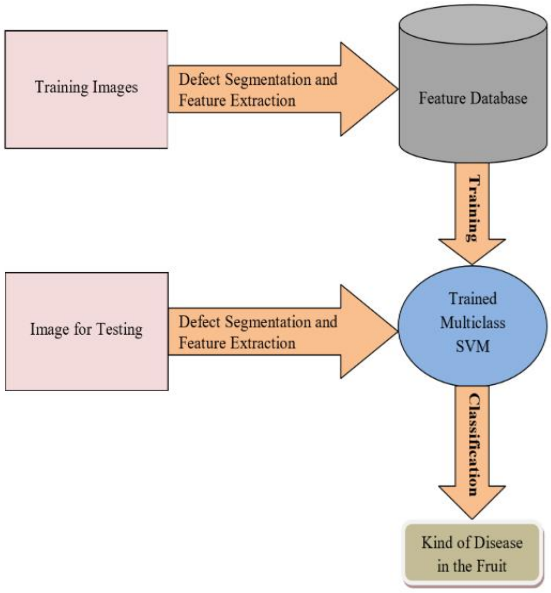



Fig. 2 (k-means clustering)

Finally, fruit and leaf diseases are classified using a Multi-class Support Vector Machine. The significance of using clustering technique for the disease segmentation and Multi-class Support Vector Machine as a classifier for the automatic classification of fruit and leaf diseases is shown. In order to validate the proposed approach, we have considered three types of the diseases in apple fruit; apple blotch, apple rot and apple scab and three types of diseases in leaves; Alternaria Leaf Spot/Blight , Marsonina Coronaria and Apple Scab . The experimental result validates that the proposed approach can significantly achieve accurate detection and automatic identification of fruit diseases. [19]

The system applies algorithm to derive vital parameters related to the properties of the leaf and fruit.

### *A. Capturing various images:*

The proposed system aims at capturing images of fruits and leaves and then processing the images with complex background, varying lightening conditions and clicked from various distances. This makes the system more dynamic to work under various climatic conditions.

### *B. Extraction Using K- Means Clustering*

After removal of background, k-means clustering algorithm is implemented on the background removed images further to

classify it into different clusters. The proposed algorithm gives the better results which was unable to compute using k-means algorithm alone. [20]

*C. Segmentation*

It is a convenient and effective method for detecting foreground objects in images with stationary background. Background subtraction is generally used class of techniques for segmenting objects of interest in a scene. This task has been widely studied in the literature. Specular reflections, background clutter, shading and shadows are the major factors that affect the efficiency of the system. Therefore, in order to reduce the scene complexity, it might be interesting to perform image segmentation focusing on the object description only. [21]

*D. Classification through Support Vector Machine (SVM)*

Support vector machine is a new promising, non-parametric and non-linear statistical classification technique. It finds its wide range of application in medical diagnostic, optical character recognition, disease prediction, disease detection and many more. During training of SVM model, kernel function is passed as a parameter which helps to do certain calculation faster and decreases time. If one has few points in a high dimensional space, linear SVM is a better option. Classification of Support vector machinery (SVM) results in statistical learning theories. [22]

*E. Naive Bayes classification*

These classify data based on probability and Bayes theorem and are also called probabilistic classifier. One of the important application is automatic medical diagnosis. Naive Bayes approach provides probability of a real time data pertaining to a particular class. The highest probability class is considered as the most probable class. It is scalable and fast algorithm used for both multiclass and binary classification which can be trained on a small dataset. [23]

*F. Decision Tree*

This approach is used in statistics, data mining and machine learning. Decision tree categorize data points. Data is passed to different nodes of the tree based on some conditions and data is assigned to respective nodes. Top node is called root whereas nodes which cannot be classified are called leaf nodes. Further at each node probability of each data point is calculated and it gives an impression of its occurrence at a particular node. [24]

## IV. CONCLUSION:

In this paper, several research papers are studied with the purpose of finding the efficient solution of disease detection in leaves and fruits of apples. From the review, a more generalized work flow is derived to follow a similar work. According to this flow, features are extracted from the image of diseased apples. A classification algorithm is used to instruct the machine for identification of disease. Such a trained machine is used to find the name of disease in a new image uploaded by the farmer. This kind of system will help farmers to improve the quality and quantity of crop.[25]

## V. REFERENCES:

[1]. Dubey, S. R., & Jalal, A. S. (2012b)"Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns" IEEE, no. July, 2016.

[2]. Dubey, S. R. (2012). Automatic Recognition of Fruits and Vegetables and Detection of Fruit Diseases. Master's theses, GLA University Mathura, India IEEE, no. July, 2016.

[3]. S.R. Dubey," Apple disease classification using color, texture and shape features from images", IEEE, no. July, 2016.

[4]. "Apple Scab and Marsonina Coronaria Diseases Detection in Apple Leaves Using Machine Learning", Swati Singh and Sheifali Gupta Volume 118 No. 18 2018, 1151-1166].

[5] Konstantinos P. Ferentinos, "Deep learning models for plant disease detection and diagnosis Computers and Electronics in Agriculture", 145 IEEE, (2018) 311–318.

[6]. M. Jhuria, A. Kumar, and R. Borse."Image processing for smart farming: Detection of disease and fruit grading", In Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013) IEEE, December 2013.

[7]. S. Sankarana, A. Mishraa, R. Ehsania, and C. Davisb." Review of advanced techniques for detecting plant diseases",Computers and Electronics in Agriculture, Elsevier, 72(1):1–13, June 2010.

[8]. D. Deshmankar, A. Awate, G. Amrutkar, U. Bagul, and S. Sonavane."Fruit disease detection using color, texture analysis and ann", In 2015 International Conference on Green Computing and Internet of Things(ICGCIoT), pages 970–975. IEEE, October 2015.

[9]. Padaliya Dharm, Pandya Parth, Patel Bhavik,Salkiya Vatsal,Solanki Yuvraj,"A-Review-of-Apple-Diseases-Detection-and-Classification" International Journal of Engineering Research & Technology (IJERT)Vol. 8 Issue 04, April-2019

[10]. "Classification of Healthy and Rot Leaves of Apple Using Gradient Boosting and Support Vector Classifier ",(IJITEE) ISSN: 2278-3075, Volume-8 Issue-12, October 2019.

[11]. Miss.Kambale , Anuradha Manik ,Dr.Mrs.Chougule "Grading Of Apple Fruit Disease", ISSN: 22779655 (I2OR), Publication Impact Factor: 3.785

[12]. Karen Simonyan, Andrew Zisserman,"Very deep convolutional networks for large-scale image recognition, Computer Vision and Pattern Recognition ",(2014).

[13]. Rumpf T., Mahlein A.K., Steiner U., Oerke E.C., Dehne H.W., Pl¨umer L."Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance", Computers and Electronics in Agriculture 74(1) (2010), 91–99.

[14]. Haloi Mrinal,"Improved micro aneurysm detection using deep neural networks, Computer Vision and Pattern Recognition", (2015).

[15] K. A. Garrett, S. P. Dendy, E. E. Frank, M. N. Rouse, and S. E. Travers, "Climate change effects on plant disease: genomes to ecosystems," Annual Review of Phytopathology, vol.44, pp.489–509,2006.

[16]. S.M.Coakley, H.Scherm,and S.Chakraborty,"Climate change and plant disease management,"Annual Review of Phytopathology, vol.37, no.1, pp.399–426,1999.

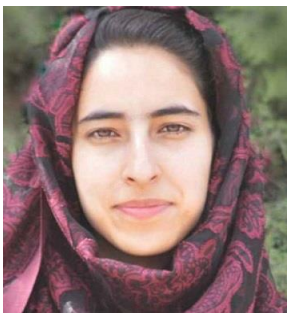
[17]. S. Chakraborty, A. V. Tiedemann, and P. S. Teng, "Climate change: potential impact on plant diseases," Environmental Pollution, vol.108 no.3,pp.317–326,2000.

[18]. S.Sankarana, A.Mishraa, R. Ehsania, and C. Davisb. "Review of advanced techniques for detecting plant diseases Computers and Electronics in Agriculture", Elsevier, 72(1):1–13, June 2010.

[19]. Anandha krishnan MG Joel Hanson1, Annette Joy2, Jerin Francis3, "Plant Leaf Disease Detection using Deep Learning and Convolutional Neural Network", IJESC Volume 7.

[20]. X. Tang, K. Liu, J. Cui, F. Wen, and X. Wang. "Intentsearch: Capturing user intention for one-click internet image search. IEEE

Transactions on Pattern Analysis and Machine Intelligence", 34(7):1342–1353, July 2012.
[21]. Nursuriati Jamil, Azlinah Mohamed, Syazwani Abdullah," Automated Grading of Palm Oil Fresh Fruit Bunches (FFB) using Neuro-Fuzzy Technique" 2009 International Conference of Soft Computing and Pattern Recognition, 4- 7 Dec. 2009, Pp. 245-249.
[22]. F.Mokhtarian and S. Abbasi,"Matching shapes with self-intersection application to leaf classification, " IEEE Trans. Image Process., 13 (5), pp. 653-661 , 2004.
[23]. Y. Tian, E. Li, L. Yang, and Z. Liang, "An image processing method for green apple lesion detection in natural environment based on GA-BPNN and SVM," in *2018 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 1210–1215, Changchun, China, 2018.
[24]. Sutton, T. B., Aldwinckle, H. S., Agnello, A. M., and Walgenbach, J. F., eds. 2014. Compendium of Apple and Pear Diseases, 2nd Edition. APS Press, St. Paul 224 pp.
[25]. Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,"Hindawi Publishing Corporation, Computational Intelligence and Neuroscience Volume 2016, Article ID 3289801, 11 pages http://dx.doi.org/10.1155/2016/3289801.



Worked on web designing and development. I have also made projects on artificial intelligence like Home Automation. I have Published paper on Smart Dust Technology in 2nd National Conference Rimt University On Recent Advances In Multidisciplinary Research.