ISSN 2278-3091
Volume 9 No.2, March - April 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse211922020.pdf
https://doi.org/10.30534/ijatcse/2020/211922020

# Orange Fruit Disease Classification using Deep Learning Approach

**Rupali Saha[1], Dr. Shubhangi Neware[2]**
[1]M Tech student, Department of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India, saharupali01@gmail.com
[2]Assistant Professor, Department of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India, newares@rknec.edu

## ABSTRACT

Orange fruit disease prediction carried out in farms is not using technology to the fullest for the production of good quality oranges. This leads to financial losses and reduction of profit to farmers. Classification and identification of diseases can be done using image processing at a low cost. In the proposed work, orange fruit images are given as input; segmentation uses color threshold to segment the images. In feature extraction, eight features which are specific to orange fruit are extracted. Then deep learning approach is used to recognize the orange fruit diseases at the initial phase of production. It can give proper treatment to the orange tree so that the cultivation of good quality orange fruit can be yielded. The overall accuracy of the system achieved is 93.21%. Hence, this paper aims to automate the process of orange fruit classification and identification, thereby maximizing economic benefits and improving the quality of life of farmers.

**Key words:** Orange disease, Image processing, Color threshold segmentation, deep learning

## 1. INTRODUCTION

Orange fruit is one of the most common citrus fruit species grown in India. India ranks second in the production of oranges in the world, due to which disease classification of oranges is necessary at an early stage of production to help the farmers to yield maximum profit. There are various varieties of oranges - Mandarin, Bitter, Bergamot, and Trifoliate. Mandarian oranges are cultivated in most of the world and are mainly affected by the fungal and bacterial infections which are classified into Brown Rot, Citrus Canker and Melanose diseases. This can be identified by the deep learning approach which is a subset of machine learning for artificial intelligence (AI). It has networks which are able to train the system that may be supervised or unsupervised. Brown Rot is a fungal pathogen that occurs in mature plants in wet condition and can be cured by copper fungicide. Melanose occurs in young plants and can be cured by liquid copper fungicide. Citrus Canker is a highly contagious bacterial infection which causes yellow halo-like lesion on the orange fruit; copper-based bactericides can provide a barrier for the disease. Hence biologists need to analyze and extract the specific contents for further classification. And here the role of image processing plays an important role [1]. A proper remedy i.e. identification of the disease is required. It is done by the proposed system consisting of two crucial stages 1) training stage and 2) testing stage. The two stages consist of three common steps – a) preprocessing. Here, noise from the image is removed using average filtering. b) Segmentation. Here, color threshold segmentation is used to segment the image. c) Feature extraction. After segmentation, the data is passed on to feature extraction unit. Here, seven features are extracted of the orange fruit. Further, database training is done using orange images and stored in the database. Images for testing are given to the classification unit, which uses convolutional neural network (CNN) to identify the diseases.

**Figure 1:** Sample images of dataset (a) brown rot (b) Citrus Canker (c) Melanose (d) Healthy orange

## 2. RELATED WORK

In cultivation of oranges, diseases may cause huge financial losses to farmers. To reduce these losses, many systems have been proposed in the past and many systems are under process.

In paper [2] apple fruit grading is performed by clustering fruit into premium, regular and rejected apple using K-means and ANN is used for the classification of the tested apple fruit.

In paper [3] deep learning approach has been used with different layers to detect mulberry leaf disease. First layer is the convolution layer which resizes the image into 3*3 matrixes. The second layer i.e. pooling layer divides the image into four non-overlapping segments to decrease over-fitting changes. In third layer i.e. activation layer, rectified linear unit (ReLU) (mathematically defined as y = max (0, x)), convolution function is used to reject negative values. In the fourth layer i.e. the fully connected layer, features are compared to get the desired results.

In paper [4], fruits are identified and their diseases are analyzed using K-means for clustering and multi-class support vector machine for classification (with an accuracy of 92.17%).

In paper [5], orange fruit quality analysis is done using Gray-Level Co-Occurrence Matrix (GLCM) for feature extraction; classification is done using threshold based segmentation and SVM.

In paper [6], there are three steps namely contrast stretching, noise filtering, histogram modification, which have been used for preprocessing. Fruit disease detection for Big Data is classified using SVN.

In paper [7], Machine learning and Fuzzy logic is used for disease classification. Grading of orange uses image enhancement and lab color transformation (i.e. L*a*b* color space is used to reduce computation time). K-means is used for segmentation; and SVM is used for classification which has obtained 90% accuracy.

In paper [8] first image is converted from RGB to HSV. Histogram of Oriented Gradients (HOG) is used to extract component descriptors, HU moments, Haralick texture, Color Histogram which is given to Random Forest classifier to classify the plant disease.

In paper [9] they have classified and detected apple fruit's three diseases with an accuracy of 95%. Here four features are extracted which is given to Learning Vector Quantization Neural Network (LVQNN). It can be used as binary or multi class classifier.

In paper [10] hybrid approach for apple fruit disease detection is used. K-means clustering is used for segmentation. Color and texture features are given to the Random Forest classifier which builds decision tree to increases predictive power of the algorithm. It also helps prevent over fitting issue.

In paper [11] four features - color, morphology, texture, structure of hole on the fruit (SURF) is extracted to identify the apple, grape, and pomegranate diseases. They are classified using ANN.

In paper [12] they have identified leaf disease and graded the leaf using image. This image is converted into binarization; and K-means is used for segmentation of different leafs, for identification of leaf disease. They have extracted eleven features classified using ANN.

In paper [13] they have identified two diseases of apple and grape fruit graded for consumption. Artificial neuron network (**ANN**) is a computational model which uses the structure and functions of biological neural networks. ANNs are considered nonlinear statistical data modeling tools where the complex relationships between inputs and outputs are modeled or patterns are found for classification of results.

In paper [14] 90% accuracy in diagnosis of pomegranate plant disease using K-means for clustering and ANN for classification has been achieved.

In paper [15] accuracy of 85.52% is acquired in identification of five cotton leaf diseases namely - Bacterial leaf blight of cotton, Alternaria, Fusarium wilt, Cotton leaf curl, Myrothecium, which is identified using Back propogation Neural Network.

In paper [16] they have detected three diseases of apple fruit. K-means is used for classification. Four features are identified - Global Color Histogram (GCH), Color Coherence Vector (CCV), Local Binary Pattern (LBP) and Complete Local Binary Pattern (CLBP). It is then given to Multi class Support Vector Machine (MSVM) to obtain 93% accuracy.

## 3. PROPOSED SYSTEM

The proposed system starts with the images of oranges given to the preprocessing unit, segmentation unit, feature extraction unit, and training. Classification, then, gives the result. The proposed system is implemented using Anaconda Navigator in python. The dataset consist of 68 images in total out of which 9 images are of Brown Rot, 19 images are of Citrus Canker, 20 images are of Melanose diseases and 20 images of healthy oranges.

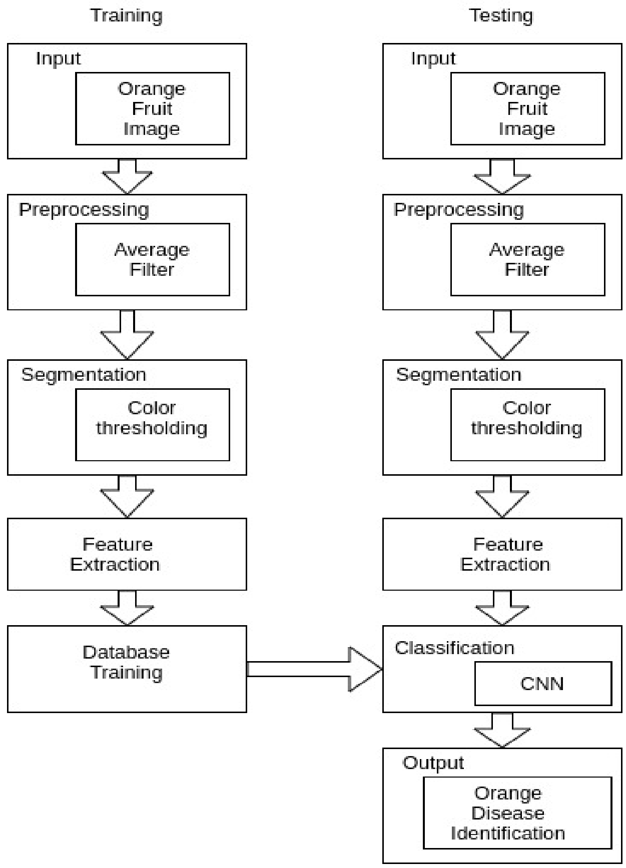


**Figure 2:** Architecture of the proposed system

### 3.1. PREPROCESSING

To make the system work accurately, the first step is to remove the noise from the image. Here, average filtering has been used as a method to 'smoothen' the images by reducing the intensity difference between neighbouring pixels. The average filter works by moving the image pixel by pixel, replacing each value with the average value of its neighbouring pixels. After this, the image is plotted in a graph of dimension 14*14 .The dimensions of the matrix in the graph represent each pixel of the image.

### 3.2. SEGMENTATION

In the previous work done they had used the gray scale images. But in our proposed work, we have used color images to get better results. In segmentation, unused part of the image is removed i.e. only red and green color should be present and red should be more than green.

### 3.3. FEATURE EXTRACTION

The proposed system, we have used texture and color features to obtain efficient and accurate results. The eight features extracted are i) vegetation index ii) normalized difference water index iii) normalized difference vegetation index iv) green difference vegetation index v) enhanced vegetation index vi) difference vegetation index vii) infrared percentage vegetation index and viii) chlorophyll index.

i. Vegetation Index (VI)

Vegetation Index (VI) may be a spectral transformation of 2 or a lot of bands designed to reinforce the contribution of vegetation properties and permit reliable abstraction and temporal inter-comparisons of terrestrial chemical action activity and cover structural variations.

ii. Normalized Difference Water Index (NDWI)

Normalized Difference Water Index (NDWI) is to mirror wet content in plants, to be determined by analogy with NDVI as:

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR}$$

NIR - near cardinal with vision in the range of 0.841 - 0.876 nm
SWIR - a part of the range with vision in the range of 1.628-1.652 nm

iii. Normalized Difference Vegetation Index (NDVI)

The normalized difference vegetation index (NDVI) is a normalized ratio of the NIR and red bands.

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

NIR - reflection in the near-infrared spectrum
RED - reflection in the red band of color

iv. Green Difference Vegetation Index (GDVI)

Green Difference Vegetation Index (GDVI) is a measure of the healthy green vegetation and the common range of GDVI is from 0.2 to 0.8.

$$GDVI = NIR - Green$$

v. Enhanced Vegetation Index (EVI)

Enhanced Vegetation Index (EVI) values ought to vary from zero to one for vegetation pixels. Bright options like clouds and white buildings, at the side of dark options like water, may result in abnormal picture element values in associate degree EVI image.

$$EVI = 205 * \frac{(NIR - RED)}{(NIR + 6 * RED) - (7.5 * BLUE + 1)}$$

vi. Difference Vegetation Index (DVI)

Difference Vegetation Index (DVI) distinguishes between soil and vegetation, however it doesn't account for the distinction between coefficient and radiance caused by region effects or shadows.

$$DVI = NIR - RED$$

vii. Infrared Percentage Vegetation Index (IPVI)

Infrared Percentage Vegetation Index (IPVI) is functionally the same as NDVI, but it is computationally faster. Values range from 0 to 1.

$$IPVI = \frac{NIR}{NIR + RED}$$

viii. Chlorophyll Index (CI)

The chlorophyll index is employed to calculate the overall chlorophyll content of the leaves. The CIgreen and CIred-edge values are sensitive to little variations within the chlorophyll content and consistent across most species.

$$CI_{green} = \frac{\rho_{NIR}}{\rho_{green}} - 1$$

$$CI_{red-edge} = \frac{\rho_{NIR}}{\rho_{red-edge}} - 1$$

## 3. 4. CLASSIFICATION

The classification process uses the training set from which features are extracted; later deep learning (which is the subset of the machine learning) is used. The classifier convolutional neural network (**CNN**) can be applied in image processing, to perform supervised learning to identify the affected orange fruits.

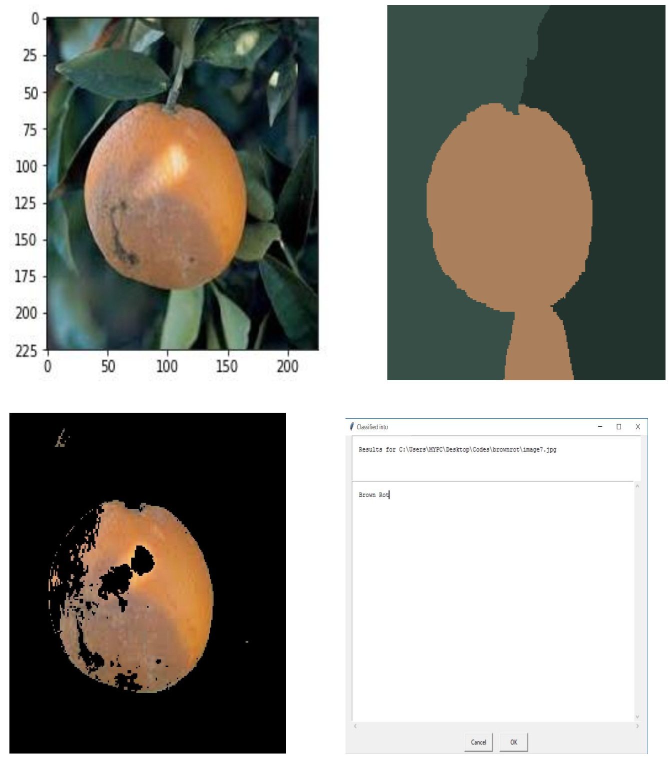


**Figure 3:** (a) Brown rot image (b) color image to RGB image (c) Segmented image (d) Orange fruit disease classified

## 4. RESULT

The proposed system classifies diseases of orange fruit and healthy oranges with the accuracy of 88.89% for the Brown rot, 84.21% for the Citrus Canker, 100% for the Melanose, 100% for the Healthy Oranges.

| Disease Name | No. of Samples | No. of Samples Correctly Identified | No. of sample Not Identified Correctly | Accuracy (in %) |
|---|---|---|---|---|
| Brown rot | 9 | 8 | 1 | 88.89 |
| Citrus Canker | 19 | 16 | 3 | 84.21 |
| Melanose | 20 | 20 | 0 | 100 |
| Healthy Orange | 20 | 20 | 0 | 100 |

**Table 1:** Table showing the result of dataset

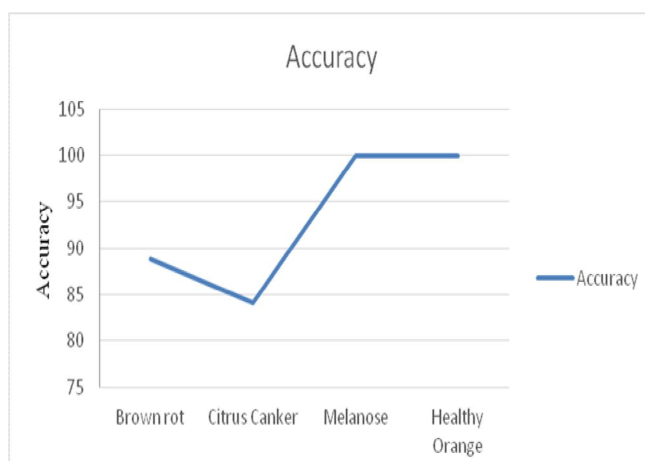


**Figure 4:** Graph showing accuracy of the proposed system

## 5. CONCLUSION

The Proposed system is used to identify unhealthy oranges and give the remedy, which will help farmers to cultivate the healthy oranges. In future, the number of diseases the system identifies could be increased; and disease severity can also be identified.

## REFERENCES


1. H. D. Gadade, and Dr. D. K. Kirang**, Machine Learning Approach towards Tomato Leaf Disease Classification**, *International Journal of Advanced Trends in Computer Science and Engineering*, Volume 9, No.1, January –February 2020 https://doi.org/10.30534/ijatcse/2020/67912020
2. S. Neware, **Fruit Grading System using k means clustering and Artificial Neural Network**, *International Journal of Advanced Trends in Computer Science and Engineering*, Volume 9, No.1, January –February 2020 https://doi.org/10.30534/ijatcse/2020/95912020
3. D. Deva Hema, S. Dey, Krishabh, and A, Saha**. Mulberry Leaf Disease Detection using Deep Learning**, *International Journal of Engineering and Advanced Technology*, Vol. 9 Issue-1, October 2019 https://doi.org/10.35940/ijeat.A1521.109119
4. S. Behera, A. Rath and P. Sethy. **Automatic Fruits Identification and Disease Analysis using Machine Learning Techniques**, *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8 Issue-6S2, April 2019
5. Kavita Komal and Dr. Sonia. **GLCM Algorithm and SVM Classification Method for Orange Fruit Quality Assessment**, *International Journal of Engineering Research & Technology*, Vol. 8 Issue 09, September-2019
6. M. T. Vasumathi and M. Kamarasan. **Fruit Disease Prediction Using Machine Learning Over Big Data**, *International Journal of Recent Technology and Engineering*, Vol. 7, Issue-6S5, April 2019
7. S. Behera, L. Jena, A. Rath and P. Sethy. **Disease Classification and Grading of Orange using Machine Learning and Fuzzy Logic**, *International Conference on Communication and Signal Processing*, April 3-5, 2018, India https://doi.org/10.1109/ICCSP.2018.8524415
8. S. Ramesh, R. Hebbar, Niveditha M., Pooja R., P. Bhat N, Shashank N, et al. **Plant Disease Detection Using Machine Learning**, *International Conference on Design Innovations for 3Cs Compute Communicate Control*, 978-1-5386-7523-6/18/$31.00 ©2018 IEEE
9. V. Jamdar and A. P. Patil. **Detection and Classification of Apple Fruit Diseases using K-means clustering and Learning Vector Quantization Neural Network**, *International Journal of Scientific Development and Research*, Vol. 2 Issue 6, 2017
10. J. Samajpati and S. D. Degadwala. **Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier**, *International Conference on Communication and Signal Processing*, April 6-8, 2016, India 978-1-5090-0396-9/16/$31.00 ©2016 IEEE
11. A. Awate, D. Deshmankar, G. Amrutkar, U. Bagul and S. Sonavane. **Fruit Disease Detection using Color, Texture Analysis and ANN**, *International Conference on Green Computing and Internet of Things*, 978-1-4673-7910-6/15/$31.00 c 2015 IEEE
12. A. Rastogi, R. Arora and S. Sharma. **Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic**, *2nd International Conference on Signal Processing and Integrated Networks*, 978-1-4799-5991-4/15/$31.00 ©2015 IEEE
13. M. Jhuria, A. Kumar and R. Borse. **Image Processing for Smart Farming: Detection of Disease and Fruit Grading,** *Second International Conference on Image Information Processing (ICIIP-*2013), 978-1-4673-6101-9/13/$31.00 ©2013 IEEE https://doi.org/10.1109/ICIIP.2013.6707647
14. M. Dhakate and A. B. Ingole. **Diagnosis of Pomegranate Plant Diseases using Neural Network**, 978-1-4673-8564-0/15/$31.00 ©2015 IEEE
15. P. R. Rothe and R. V. Kshirsagar**. Cotton Leaf Disease Identification using Pattern Recognition Techniques**, *International Conference on Pervasive Computing*, 978-1-4799-6272-3/15/$31.00(c)2015 IEEE
16. S. R. Dubey and A. S. Jalal. **Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns**, *Third International Conference on Computer and Communication Technology*, 978-0-7695-4872-2/12,© 2012 IEEE