JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 728 - 735
https://publishoa.com
ISSN: 1309-3452

Plant Leaf Disease Classification and Prediction Using a Customized Deep Transfer Learning Model

Chandrika Bhargavi Achanta[1], Kavuru Devi Keerthi[2], Sujatha Kamepalli[3]
[1]VFSTR Deemed to be University, Guntur, A.P., India.
[2]VFSTR Deemed to be University, Guntur, A.P., India.
[3]VFSTR Deemed to be University, Guntur, A.P., India.
sengottuvelan@gmail.com
(Corresponding Author: S.Joseph Gabriel)

ABSTRACT

There is a significant productivity, financial damage due to plant diseases, and diminished overall quality of farm products. Detecting plant diseases has become more important in the surveillance of vast fields of crops in the modern-day. When it comes to disease management, farmers have difficulty transitioning from one strategy to another. The conventional method for detecting and identifying plant diseases is professional naked-eye inspection. This research examines the necessity for a simple technique for detecting plant leaf disease that would aid agricultural innovations. Early knowledge of crop health and disease detection may facilitate effective monitoring tactics. Crop yields will rise as a result of this method. In addition, the advantages and drawbacks of each of these prospective approaches are discussed in this study. Image capture, image analysis, extraction of features, and categorization based on neural networks are all part of the process. We get the best result to help the farmers through the processed methodology by implementing this model. The resulted accuracy of the implemented model is 81.09492659568787%. The proposed work enhances the farming culture to predict certain diseases and get a good yield of crops.

Keywords: Plant leaf Diseases, Deep Learning, VGG-19 Model, Crop Yield.

1 Introduction

Plants get affected by various diseases due to pathogens and environmental conditions. The disease occurrence varies from season to season. To get out of these crop diseases, farmers use different types of pesticides to control the disruption caused by the pathogen and provide security for the plant to get much yield as per their farming investment. If any plant of a particular crop is subjected to a disease outbreak, then the leaves of the plant change drastically with specific features. It is due to the infection caused by that disease, resulting in a decrease in crop yield and loss of crops. Incredibly less developed countries where admittance to disease control methods is minimum should face hunger and starvation. In this proposed algorithm, we use a transfer learning model to process the image input. Transfer learning is a machine learning technique where we use the knowledge obtained from the pre-trained model as the entry point for the model of a new task. It results in optimization that allows rapid progress when modeling the new task. It ensures higher performance even for a large amount of data. The target model obtained from the transfer learning model is highly efficient with reasonable accuracy. Out of various transfer learning models, we used VGG19 to recognize the disease by taking the plant leaves and processing those leaves images through the pre-trained model.

Fig. 1 Infected leaf of Apple Plant

2 Literature Review

Several studies describe how to identify and treat the disorders. Techniques illustrating how to put them into practice, as seen and being the subject of this discussion, the possibility of finding diseased cotton leaves. (Anand.H.Kulkarni, 2012). A method in which the color and spot characteristics may be retrieved backward propagation of a self-organizing feature map a network of neurons detection of a diseased area on the leaves of a plant and texture traits are used to classify plant leaf diseases. An algorithm was devised by (S. Arivazhagan, R. Newlin Shebiah, S. Ananthi & Varthi, 2013) to process the data, transforming the supplied color. The red, green, and blue components of an RGB picture are green pixels disguised after being formed. Shape and form texturing characteristics are considered from the data. In order to classify objects, the Minimum Distance is used. Vector machines for criterion evaluation and support (SVMs). (Al Bashish et al., 2010), they have developed a framework for the Identification and Classification of Plant Diseases of the Leaf and Stem. The K-Means algorithm uses to segment and transforms RGB pictures into HIS Tonal system. Color and texture characteristics are then calculated. A statistical neural network classifier. Classification is used for categorization purposes. Image segmentation has been hypothesized by (Nandini Kakran & Pratik Singh, 2019) for crop disease detection. The three-part technique is as follows: A device-independent color format, RGB, was used to store the leaf pictures at first. To separate the photos, they had to be resized and converted to the CIELAB color space independent of hardware. Region-based extraction of the contaminated area is the second method. (K. Sujatha et al, 2019) The use of segmentation was made. (S. R. Bandaru et al, 2019) An essential part of K-means clustering is A method used during the segmentation stage. Twelfth extract characteristics are based on color, shape, and texture. Typically used to describe a geographic area. As a result, The co-occurrence matrix of characteristics, wavelet, and grey levels. It was necessary to make use of certain strategies. Classifying apples using a support vector regression (SVR) algorithm disease of the leaves. Agricultural plant disease has been hypothesized by (Sanjay B. Dhaygude, 2013); a prototype software approach for detecting rice disease based on diseased photos of diverse rice plants is described in Identification Using Pattern Recognition Techniques. Photos of diseased rice plants are taken using a digital camera, and sick sections of the plants are identified using image growth and image segmentation algorithms. After that, a neural network is used to classify the affected leaf tissue. Cotton Leaf Spot Diseases Detection Using Feature Selection with Skew Divergence Approach has been suggested by (Dheeb AI Bashish, Malik Braik, 2010). In this study, the upgraded PSO feature selection method uses the skew divergence method and extracts Edge, CYMK color, GA, Color, and Texture variations features. The collected features were analyzed using SVM, BPN, and Fuzzy with Edge Selection & Classification. (Yan-Cheng Zhang, Han-Ping Mao, Bo hu, 2007), A panel of experts analyzed the images on the remote server. Computer vision methods are used to identify and classify damaged areas in a picture. The disease-affected lesions are segmented using a straightforward color difference technique. (K. Sujatha et al, 2021) With the use of a mobile phone notification, the expert may assess the study's findings and offer comments to the farmers. An image recognition system capable of identifying crop diseases is the purpose of this study. The scanned color picture of a diseased leaf is the first step in image processing. These photos are segmented using a mathematical morphology technique. These attributes were then utilized in conjunction with a classification approach called the membership function to separate the three categories of illnesses. (Haiguang Wang, Guanlin Li, Zhanhong Ma, 2012a) cotton leaf disease was identified and classified using image processing and machine learning approaches. The poll results on background removal and segmentation approaches were also brought up throughout the discussion. A translation from RGB to HSV color space for background removal was discovered during this research. In addition, we discovered that thresholding produces superior results compared to other methods of removing background noise. We used thresholding to the masked image after performing color segmentation by masking green pixels in the background removed image to produce a binary picture. This segmentation process is important in extracting disease-specific characteristics. For illness categorization, we observed that SVM performed well in accuracy. A total of three of our recommended processes have been implemented: Image Acquisition, Image Preprocessing, and Image Segmentation. Our proposed work has five key parts. (H. AI-Hiary, S. Bani-Ahmad, M. Reyalat, 2011) After taking a picture, the collected pictures are initially enhanced using Image Edge detection and segmentation algorithms. Image segmentation is conducted out using the R, G, and B Color Feature images (disease spots). An image feature extraction process is then used to identify disease areas and implement a pest management strategy based on these attributes. Cotton leaf spot, cotton leaf color segmentation, and image segmentation based on edge detection are included in this study's analysis and categorization of illness. (M. Meunkaewjinda, P. Kumsawat, K. Attakitmongcol, 2008) We propose and test a new software method for detecting and classifying plant illnesses using Image Processing. Few agricultural professionals in rural India can examine crop photographs and provide advice. Farmers often get expert solutions to their inquiries too late because of the delay in expert responses. Image processing algorithms based on color, texture, and shape will be developed in this article to identify crop illnesses or other issues that may impact crops from photos and provide farmers with the most up-to-date information through SMS. As a result, more efficient use of chemicals will lead to increased production and a better-quality output due to these technologies. (Lili li, Shujuan zhang, 2021), Using deep learning and advanced imaging techniques, the disadvantages in plant disease recognition get removed, and the efficiency of the proposed algorithm is also increased by using a particular methodology in processing the application. Certain visualization techniques are used to organize deep learning architectures. A roll of metrics is used for training and validation. Mean Average Precision (MAP) is an indicator for evaluating the framework. (Haiguang Wang, Guanlin Li, Zhanhong Ma, 2012b) In this proposed technique, Principal component analysis (PCA) and Backpropagation networks (BP) are used to predict two kinds of wheat diseases. Generalized Regression Networks GRNNs and Probabilistic Neural Networks PNNs are used for classification to identify the grape diseases. The fitting and prediction accuracy of the class labels are identified via these processing algorithms. (Song kai, Liu zhikun, Su hang, 2011) Based on image processing and analysis for maize crop disease prediction, image recognition of corn leaf. Texture features of corn disease and segmentation of disease spots are used to recognize the disease. (K. Sujatha et al, 2020) The backpropagation neural network (BP) mechanism is also implemented to coordinate disease detection input/output relations. A spatial gray matrix is used for fast identification of the maize disease. (Mokhled S, 2013) image analysis and classification techniques for spotted disease detection of Olive leaves are used for the research. Various images of leaves were collected from the field and taken as samples. C- means clustering is used to detect the defect. The severity of the disease is determined via the classification of leaf areas. As a result, 86% accuracy is obtained through the processed algorithm. (Argenti et al., 1990) An algorithm for fast analysis is proposed for calculating the parameters of matrices. A supervised learning mechanism is used for reading the input data. Statistical features get extracted through classification and segmentation.

3 Objectives

The primary contributions of this research are:

1. We understand the features of infected and uninfected leaves of different fruits and vegetable crops.
2. We explore the existing machine learning and deep learning models used in plant leaf disease detection.
3. We are developing deep transfer learning models to detect plant leaf diseases and customizing the model to work with different class labels.
4. We are testing the model with randomly selected plant leaf images to predict whether the given leaf is infected or uninfected, finding the disease name.

4 Methodology

The basic steps involved in plant leaf disease detection are given below.

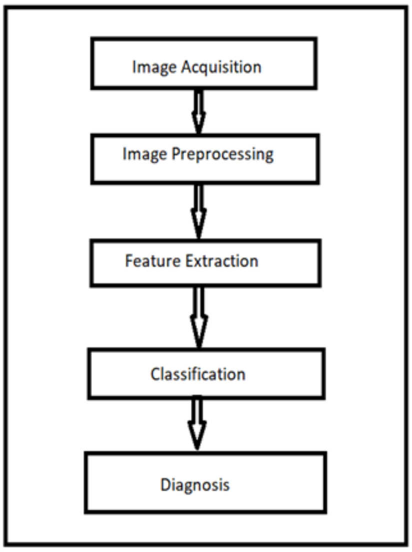

Fig. 2 Basic Steps for Disease Detection Algorithm

4.1 Acquisition of images

High-quality RGB color photographs are taken using a digital camera with the necessary resolution. The project explicitly dictates how an image database is built. The classifier, which determines the algorithm's resilience, is made more efficient by the picture database.

4.2 Processing and segmentation of images

During the pre-processing stage, background and noise are removed from the picture data, and unwanted distortions are suppressed. It improves picture processing and analysis capabilities.
HIS and CIELAB color spaces are used to transform the RGB color picture. When HSI and CIELAB were developed as color-independent space models, they were also based on human perception.
Segmentation is the first step in locating the area that is contaminated.
K-means clustering and edge detection algorithms are often used for segmentation.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 728 - 735
https://publishoa.com
ISSN: 1309-3452

A variety of characteristics are used to define the contaminated area after segmentation. For describing a location, color, texture, and form are often employed.
Color characteristics are essential for determining the visual surroundings, identifying objects, and conveying information. An object's texture is one of its most critical identifying characteristics. Image retrieval benefits from its use as a robust regional descriptor.
The texture is described, by contrast, as homogeneity, dissimilarity, energy, and entropy. An image's shape is one of the most fundamental properties that may be used to describe its content.

4.3 Classification

At this point, no other diseases may be detected. Classifiers such as the Support Vector Machine and the Artificial Neural Network are often employed.
According to specified traits and allocating each illness to one of the preset groups, it is identifying a rule.

The transfer learning model VGG-19 was customized and used in plant leaf disease prediction. The following diagram depicts the basic architecture of the VGG-19 model.
The overall architecture of our research work is represented in the following flowchart.

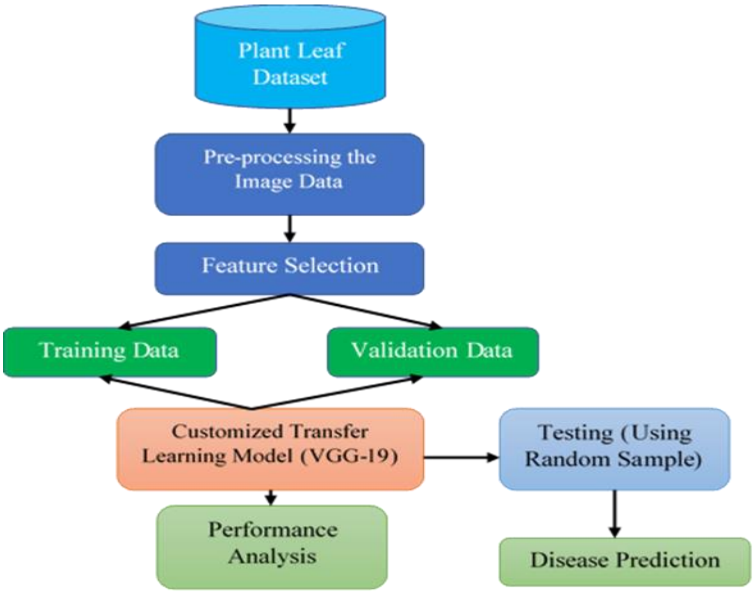

Fig.3 Overall Architecture of Proposed Model

4.4 VGG-19 Model

VGG stands for Visual Geometry Group. As the name VGG19 indicates, it consists of 19 layers. Out of those 19 layers, 16 are convolution layers, 3 fully connected layers, 5 Max pool layers, and 1 Softmax layer. It is mostly used for transfer learning mechanisms. It is also named a good classification architecture for processing the input data.

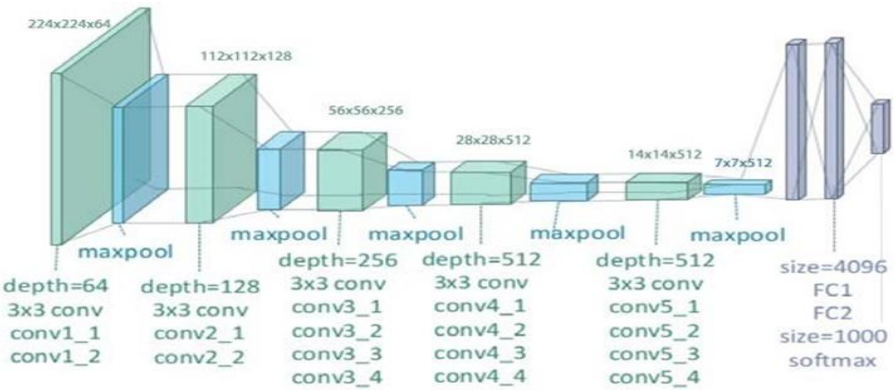

Fig. 4 VGG19 Architecture Diagram

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 728 - 735
https://publishoa.com
ISSN: 1309-3452

Source: **https://www.researchgate.net/figure/llustration-of-the-network-architecture-**of-VGG-19-model-conv-means-convolution-FC-means_fig2_325137356

5 Experimental Results

This section presents the dataset description and experimental results obtained from our implementation.

5.1 Dataset Description

We analyzed various images of the leaves of the plant. We rescaled all the images of leaves by using image processing and segmentation. The dataset consists of 38 class labels categorized into trained and validation data; as per the features and texture of the image input sent through the class labels, the plant's disease gets detected. We run all our experiments across a range of train-test set splits, namely 80–20 (80% of the whole dataset used for training and 20% for testing). The following diagram shows the different class labels considered for experimentation.

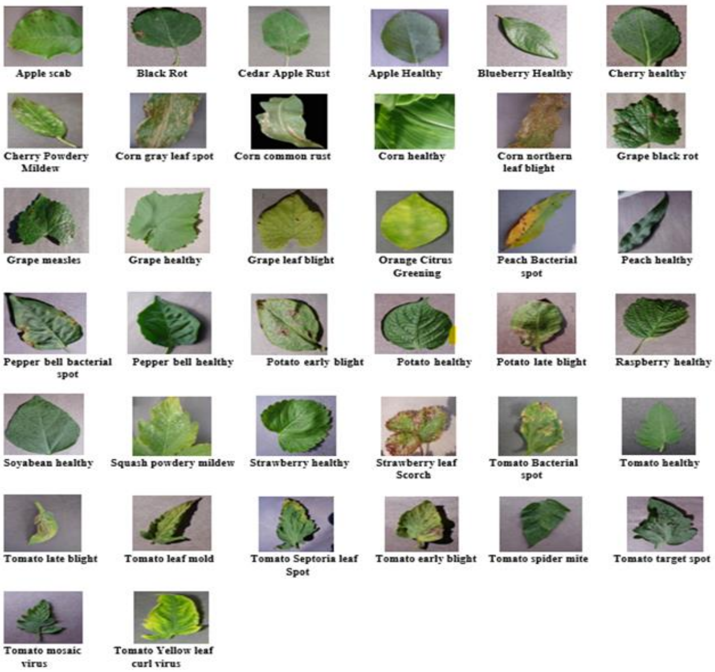

Fig. 5 Samples of Different Class Labels in the Dataset

5.2 Experimental setup

The system is configured with Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz-1.99GHz and 16GB Random Access Memory (RAM). It has a 64bit operating system with X64 based processor. Here we used the Google Collaboratory platform to execute the code.

5.3 Results and Discussions

The results obtained from the implementation of the VGG-19 model are represented in this section. We obtained 81.1% accuracy in plant leaf disease classification and prediction. The following snapshot gives the code implemented to print the model's accuracy.

```
acc = model.evaluate_generator(val)[1]

print(f"The accuracy of your model is = {acc*100} %")
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: UserWarning: `Model.evaluate_generator` is deprecated and will be removed in a future
  """Entry point for launching an IPython kernel.
The accuracy of your model is = 81.09492659568787 %
```

Fig. 6 Accuracy of the Model

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 728 - 735
https://publishoa.com
ISSN: 1309-3452

The following figures show the comparison of training accuracy and validation accuracy.

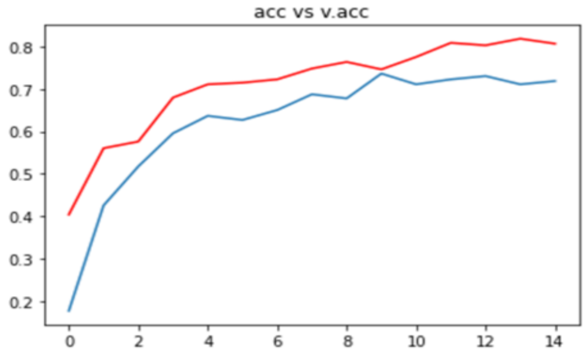

Fig. 7 Comparison of Training Accuracy and Validation Accuracy

The following figures show the comparison of training loss and validation loss.

Fig. 8 Comparision of Training Loss and Validation loss

We tested the customized VGG-19 model by giving a random plant leaf image to predict whether it is infected or not, if it is infected, and displaying the name of the disease.

Fig. 9 Testing with a Random Image (Tomato_Yellow_Leaf_Curl

The customized VGG-19 model efficiently predicts different types of plant leaf diseases. The model can predict 38 types of diseases from the plant leaves of different types of fruits and vegetables. The model accuracy is 81.1%. In the future, we will improve the accuracy of the model.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 728 - 735
https://publishoa.com
ISSN: 1309-3452

6 CONCLUSION

The central theme of this research is to predict the disease of plant leaves caused by different types of pathogens. In this approach, we used a transfer learning mechanism that initiates the usage of knowledge obtained from the previously implemented model. We took the dataset as input which consists of various images of leaves of certain fruits and vegetable plants. We take 38 class labels indicating various diseases of the train and validation datasets. Under pre-processing, the images of leaves are normalized and rescaled, and their size is customized. The transfer learning model VGG19 is implemented in 38 classes. The model's accuracy on the given dataset is 81.1%. In the future, we will improve the accuracy of the model.

References

1. Anand. H. Kulkarni, A. P. R. K. (2012). Applying Image Processing technique to detect plant diseases. International Journal of Modern Engineering Research(IJMER), 28(1–2), 54–69. https://doi.org/10.1177/0958305X16685471
2. S.Arivazhagan, R. Newlin Shebiah, S. Ananthi, & Vishnu Varthini, S. (2013). Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agricultural Engineering International: CIGR Journal, 15(1), 211–217. https://doi.org/10.1049/ip-f-2.1990.0064
3. Al Bashish, D., Braik, M., & Bani-Ahmad, S. (2010). A framework for detection and classification of plant leaf and stem diseases. Proceedings of the 2010 International Conference on Signal and Image Processing, ICSIP 2010, 113–118. https://doi.org/10.1109/ICSIP.2010.5697452
4. K. Sujatha, K. Venkata Krishna Kishore, and B. Srinivasa Rao, "Animal Breed Classification and Prediction Using Convolutional Neural Network Primates as a Case Study," in *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT) | 978-1-6654-1480-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICECCT52121.2021.9616928*, 2021, pp. 1–7.
5. Nandini Kakran, Pratik Singh, K. P. J. (2019). Crop disease detection using image segmentation. International Journal of Computer Applications, 2(4), 190–194.
6. Sanjay B. Dhaygude, N. P. K. (2013). Agricultural plant Leaf Disease Detection Using Image Processing. 2(1), 599–602.
7. Revathi, P., & Hemalatha, M. (2012). Classification of cotton leaf spot diseases using image processing edge detection techniques. International Conference on Emerging Trends in Science, Engineering and Technology, 169–173. https://doi.org/10.1109/incoset.2012.6513900
8. K. Sujatha, K. V. K. Kishore, and B. Srinivasa Rao, "Analyzing factors influencing the research progression of faculty in STEM based universities," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 5, pp. 2672–2686, 2020.
9. K. Sujatha and B. Srinivasa Rao, "Recent Applications of Machine Learning : A Survey," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 6, pp. 263–267, 2019, doi: Retrieval Number: F10510486C219 /19©BEIESP.
10. Yan-Cheng Zhang, Han-Ping Mao, Bo hu, M.-X. L. (2007). Features selection of cotton disease leaves image based on fuzzy feature selection techniques. Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, ICWAPR '07, 1, 124–129. https://doi.org/10.1109/ICWAPR.2007.4420649
11. Natarajan, V. Anantha and M.S Kumar. "Detection of disease in tomato plant using Deep Learning Techniques." International Journal of Modern Agriculture 9, no. 4 (2020): 525-540.
12. S. R. Bandaru and S. Kamepalli, "Artificial intelligence: applications, framework and concerns-BFSI," *J. Adv. Res. Dyn. Control Syst.*, vol. 11, no. 9, pp. 1–11, 2019, doi: 10.5373/JARDCS/V11I9/20192766.
13. H. AI-Hiary, S. Bani-Ahmad, M. Reyalat, M. B. and Z. Alr. (2011). Fast and Accurate Detection and Classification of Plant Diseases. International Journal of Computer Applications, 17(1), 31–38. https://doi.org/10.5120/2183-2754
14. M. Meunkaewjinda, P. Kumsawat, K. Attakitmongcol, A. S. (2008). Grape leaf disease detection from color imagery using hybrid intelligent system. Proceedings of ECTI-CON, 1, 513–516. https://doi.org/10.1109/ECTICON.2008.4600483
15. Lili li, Shujuan zhang, B. wang. (2021). Plant Disease Detection and Classification by Deep Learning - A Review. IEEE Access, 9(Ccv), 56683–56698. https://doi.org/10.1109/ACCESS.2021.3069646
16. Gampala, V., Kumar, M.S., Sushama, C. and Raj, E.F.I., 2020. Deep learning based image processing approaches for image deblurring. *Materials Today: Proceedings*.
17. Natarajan, V. Anantha, et al. "Segmentation of nuclei in histopathology images using fully convolutional deep neural architecture." 2020 International Conference on computing and information technology (ICCIT-1441). IEEE, 2020.
18. K. Sujatha and B. Srinivasa Rao, "Recent Applications of Machine Learning : A Survey," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 6, pp. 263–267, 2019, doi: Retrieval Number: F10510486C219 /19©BEIESP.

19. Mokhled S, A.-T. (2013). An Emperical Investigation of Olive Leave Spot Disease Using Auto-Cropping Segmentation and Fuzzy C-Means Classification. World Applied Science Journal, 23(9), 1207–1211. https://doi.org/10.5829/idosi.wasj.2013.23.09.1000
20. Argenti, F., Alparone, L., & Benelli, G. (1990). Fast algorithms for texture analysis using co-occurrence matrices. IEEE Proceedings, 137(6), 443–448. https://doi.org/10.1049/ip-f-2.1990.006