International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

Implementation of Ensemble Classifier for Plant Disease Detection

Shubham Gupta[1], Jasbir Singh[2], Vinod Sharma[3]

[1]M. Tech. Final Year Student, [2]Assistant Professor, [3]Professor
[1, 2, 3]Department of Computer Science and IT, University of Jammu

---

Abstract: Image processing refers to the technique for the extraction of digital information from images. These image processing techniques have been applied to various domains including medical field, remote sensing, robot vision, pattern recognition, video processing, color processing, and so on in order to solve real world problems. These approaches have also been found out to be effective in the domain of Agriculture for the prediction of crop yield, leaf disease detection, fruit disease detection, vegetable quality evaluation, etc. Due to the excessive usage of fertilizers and insecticides, the quality of plants and crops gets degraded to huge extent; and the plant diseases are generally detected by experts through naked eyes. In this paper, the authors proposed an ensemble model based on Random Forest and K-Nearest Neighbor (KNN) for the detection of plant diseases from the leaves. The experimentation has been carried out on benchmark dataset of Plant Disease Images. This dataset consists of 1000 images of Healthy Leaves and three plant diseases i.e. Brown Rust, Early Blight and Late Blight. The proposed approach works in four steps: Segmentation by K-Means Clustering, Feature Extraction by GLCM, Feature Selection by Random Forest, and Classification of plant disaese by KNN. The performance of the proposed approach has also been evaluated with traditional support vector machine (SVM) on the basis of three measures such as accuracy, precision and recall. The experimental results revealed that the proposed approach outperformed SVM.

Keywords: Ensemble Technique, Feature Extraction, Image Processing, Plant Disease, Segmentation.

1. INTRODUCTION

In many developing nations, agriculture is the spine of economy. Particularly, Agriculture plays an important role in India's economic growth. The 70% of the Indian economy depends on the agriculture to earn their livelihood. The huge loss in production would cause due to damage of harvest. In this way, the entire economic growth of a country would be affected. One of the sensitive parts of the plant is leaves. The symptoms of the disease can be seen in leaves of plant in former stages. The crop production's quantity and quality depend on the growth of plant. Thus, it is very important to detect plant leaf disease in earlier stage [1].

After disease detection, some necessary steps should be taken to prevent it from spreading to others regions of the farm. In general, the color and shape of the plant leaves are monitored by farmers for disease detection. This process requires long time expertise and lot of regular hard work. This is almost impossible for the big farms. The variations in signs, spots, color, etc can be monitored to identify several infections occurring within the different parts of the plant. The time efficient and automated diagnosis technique is the chief requirement of farming sector to enhance the productivity rate of crop. The solution to deal with several problems based on farming applications are obtained using processing methods in these days. These applications include the detection of disease, leaf, stem and fruit. A lot of researchers have focused on the measuring and recognition of Leaf disease using image processing.

In order to identify leaf infection, various image processing techniques has been introduced. Figure 1.1 illustrates the various stages that are comprised in the leaf disease detection system in which image processing is carried out.

Following are the stages involved in plant disease detection:

a. Image Acquisition
b. Preprocessing
c. Segmentation
d. Feature Extraction
e. Pattern Matching and Classification

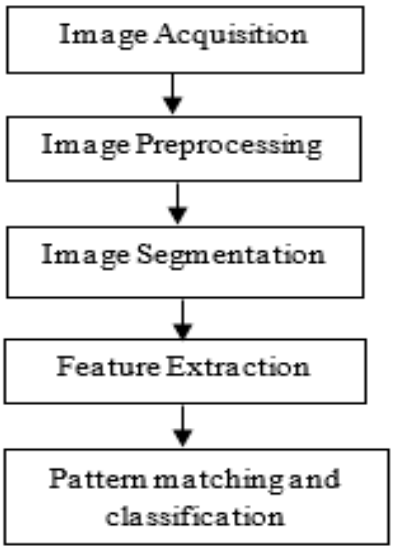

Fig. 1: Different Stages in Plant Disease Detection

A brief description of these stages is given below:

1.1 Image Acquisition:

In this process, the graphic image generation that is related to physical view or the inner structure of an entity is presented. In general, the task of restoring an image from some origin is called image acquisition. A source that is based on the hardware is employed to normally restore the image [2]. The processing of this source is done with other processes. The emerging of that processes are needed later. The fundamental state that is required always for the sequence of work flow is Image acquisition due to the necessity of an image in image processing. The image is natural which is obtained.

1.2 Preprocessing:

The major purpose of image pre-processing is to improve the picture information comprised redundant distortions or to strengthen some image features for processing. Various methodologies such as dynamic image size and form, noise filtration, image translation, image enhancement and morphological operations are utilized by preprocessing stage.

1.3 Image Segmentation:

The Image segmentation is the method in which a digital image is split into a number of segment parts. The chief objective of segmentation is that the information is extracted and the objects are recognized from the images. This process simplifies the task of picture inspection. The objects as well as bounding line of images are located using this method. The pixels that have included similar label part share the distinguished attributes so as the label is assigned to each pixel in an image. In the image segmentation, some schemes are carried out that include K-means clustering, Otsu's algorithm, thresholding and so on. These schemes are as following:

Segmentation using Boundary and spot detection algorithm: The RGB image is converted into HIS model during segmentation. The Boundary detection and spot detection have employed to identify the disease portion of the leaf [3]. The consideration of connectedness of pixels is required for detecting the boundary. After that the boundary detection algorithm is carried out.

K-means Clustering: In this method, the classification of object is done into K number of classes. This classification is based on a set of features. The reduction of sum of the squares of the distance is seen between the object and the equivalent cluster to classify the object.

Otsu Threshold Algorithm: The thresholding is carried out to generate the binary images from grey-level images [4]. Some of the pixels are set to zero that are below to threshold whereas; some of the pixels are set to one that is above from threshold in this process. The diseased leaf displays the indications of the infection by varying the color of the leaf. Therefore, it is possible to use the greenness of the leaves to detect the diseased part of the leaf. From the picture, R, G and B elements are extracted. Otsu's method is used for calculating threshold. If the green pixel brightness is below the computed threshold then masking and removal of the green pixels is done.

1.4 Feature Extraction:

Up to now, the result obtained is the region of interest. Therefore, this step is implemented to extract the features from this region of interest. The feature extraction is a procedure in which a set of values is extracted from an image. This set is known as features. These features provide information regarding the picture for more processing. The infection in the plant is recognized using several features. Feature extraction can be performed using different techniques. These methods can be used to develop a system [5]. The GLCM, color co-occurrence technique and feature extraction based on histogram are carried out.

1.5 Pattern Matching and Classification:

In this process of pattern matching, the sequence of tokens is analyzed for the subsistence of the

elements that are related to some pattern. The fundamental aim of this algorithm is that the features are compared with the index features of picture that are occurred in the database. The classification is a procedure which is utilized to identify the category of the pattern that are observed. Supervised and unsupervised are the two main types of classification. Training is needed in supervised classification [6]. The supervised classification is process in which the selection of sample pixels is performed to generate a class. Training is not required in unsupervised classification. The results are relied on the analysis of software in which sample classes are not included in the unsupervised classification. A number of algorithms are carried out for the detection of diseases of plant. Some of these algorithms are SVM, K-NN, neural networks and so on.

2. LITERATURE REVIEW

V. Singh, et.al (2015) introduced a renewed image segmentation algorithm. The main objective behind using this algorithm was to notice and classify the disorder in leaves [7]. Here, various types of classification algorithms were reviewed. These algorithms could be implemented to detect different diseases within plant leaf. Genetic algorithm was used in this work for image segmentation. This algorithm played an important role to detect diseases within plant leaf.

A. Devaraj, et.al (2019) stated that farming was not just a technique as it was the main source of food for ever growing population [8]. Almost 70% of the total population of Asian nations was dependent on agriculture for their livelihood. However, different types of diseases reduced the quality of crop. Losses to farming could be prevented by efficient disease detection. This work was aimed of developing a software approach for the computerized classification and detection of disease. The disease detection process included different phases. Infections within plants were detected using leaf images. Thus, it was advantageous to implement image process techniques to detect and classify leaf disorders in farming sector.

R. M. Prakash, et.al (2017) implemented image processing algorithms to detect leaf disorders [9]. Applying image analysis and classification algorithms to detect and classify leaf diseases was the main purpose of this work. There were four stages included in the suggested system. This work made use of K-means approach for segmenting leaf image to identify infectious regions. This work performed the task of texture feature extraction using GLCM features. In this work, the classification task had been performed by SVM classifier.

D. M. Sharath, et.al (2019) stated that plant diseases in farming had become a serious issue as these diseases caused losses in the production. This also affected the quality of the farming produce [10]. It was an extremely complex job to monitor the health of plant and identify different pant infections in manual manner. Expertise in the plant infection detection was required for this purpose. Moreover, this process was very time consuming. Therefore, plant infections were identified using image processing. There was various infection detection stages included in this process. Disease affecting plant was monitored on the basis of output achieved using these stages. In this work, the images of infectious plants were used to discuss the technique implemented for plant infection recognition.

G. K. Sandhu, et.al (2019) reviewed and summarized different plant infection methodologies with the help of image processing. In the last few years, these techniques attracted many researchers for detecting plant diseases [11]. These algorithms classified plant leaves as normal or contaminated. However, this procedure raised different issues. These issues included the computerization of the recognition system by complex pictures clicked in outside lighting and extreme environ circumstances. In this work, it was concluded that these infection recognition methodologies showed competence and accurateness. These techniques could run the system designed to detect plant leaf infections in spite of some limits. Thus, in future, more research work would be carried out in this area to improve the earlier studies.

D. Sachin, et.al (2015) analyzed that the losses in the harvest and quantity of the agricultural goods could be prevented significantly by identifying plant infections [12]. The study of the plant infections represented the study of visibly recognizable patterns within the plant. For sustainable cultivation, it was extremely important to monitor health of the plant and detect disease. However, manual monitoring of plant infection was an extremely complex task. This process was tedious and required proficiency in the plant diseases. Also, this process was very time consuming. Therefore, the plant infections were detected using image processing. These were different stages involved in plant disorder recognition. A discussion was made here for detecting leaf disorders with the help of leaf pictures. Different sorts of algorithmic approaches employed to detect plant disease were also talked about in this work.

R Anand, et.al (2016) introduced a renewed algorithmic approach of plant leaf infection detection. This algorithm carefully detected diseases [13]. In this work, image processing and artificial neural algorithms were used for detecting brinjal leaf infection. Leaf disorders were the major concern in this work. This reduced the produce of brinjal significantly. In this work, instead of whole brinjal plant, focus was on just brinjal plant leaf. Almost all disorder arose on the leaves of eggplant. K-means approach was used for segmenting while NN was used for classification in this work. The leaf disorders were detected efficiently using suggested approach on ANNs.

C. G. Dhaware, et.al (2017) stated that diseases within plant leaves were detected and identified by implementing image processing methodologies [14]. It was advantageous to detect plant leaf infections using some automated technique as it reduced a lot of hard work to observe disease in big agricultural fields. The disease symptoms were detected at early stage using this technique. Different steps were involved in the plant infection recognition and classification. In this work, a discussion was made on different methods used for image pre-processing and image segmentation algorithm for identifying plant disorders in automatic manner. This work was carried out on different plant leaf infection classification techniques applicable to classify leave diseases.

U. M. Korkut, et.al (2018) made use of image processing and machine learning approaches for identifying leaf infections in automatic manner [15]. Early and exact detection of leaf disorders contributed significantly for crop quality and production. The cost of plant diseases and usage of unnecessary pesticides could be reduced by timely detection and interference. Leaf images of different plant species were collected here. This work implemented TF (Transfer learning) technique for taking out important attributes of the pictures. Accuracy rate of 94 % was obtained by proposed model using different machine learning techniques.

P. B. Padol, et.al (2016) analyzed that infections within plant leaves were detected and classified using a popular approach called image Processing [16]. This work made use of SVM classifier for detecting and classifying disorder in the leaves of grape. Initially, this work implemented K-means approach for the segmentation to detect the infectious region. Later, the extraction of two particular attributes was performed. At last, different types of leaf infections were identified by classification algorithm. Accuracy rate of 88.89% had been attained by the introduced scheme in the classification and detection of diagnosed diseases.

R. Pawar, et.al (2017) stated that image processing was extremely important in plant infection detection or in grading of the quality of the fruit [17]. Losses in the production and the quantity of farming goods were reduced to a large extent by detecting leaf infections. The study of the plant infections represented the study of visibly recognizable patterns on the plant. For sustainable cultivation, it was extremely important to monitor health of the plant and detect disease. Manual monitoring of plant infections was an extremely complex task. In this paper, different Pomegranate plant disease detection techniques were applied using plant leaf pictures.

3. MATERIALS AND METHODOLOGY

The main aim of this work is it to detect plant diseases from the leaves. In this paper, the authors proposed an ensemble model based on Random Forest and K-Nearest Neighbor (KNN) for the detection of plant diseases from the leaves. The experimentation has been carried out on benchmark dataset of Plant Disease Images. This dataset is collected from kaggle (*https://www.kaggle.com/emmarex/plantdisease*). This dataset consists of 1000 images of *Healthy Leaves* and three plant diseases i.e. *Brown Rust*, *Early Blight* and *Late Blight*. Figure 2 represents the images of the plant disease in the dataset.

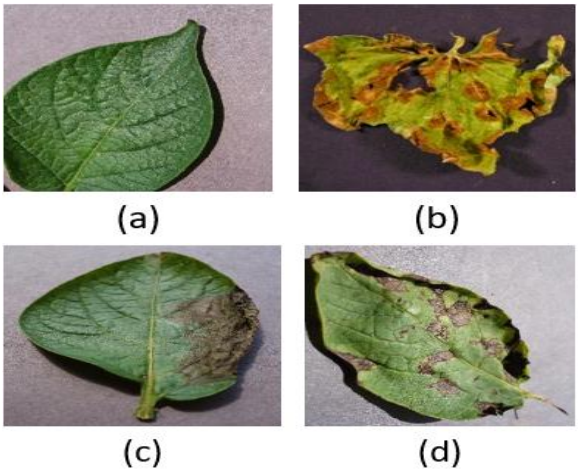

Fig. 2: (a) Normal leaf (b) Brown Rust (c) Early Blight (d) Late Blight

The proposed approach is presented in figure 3, and it works in four steps: Preprocessing, Segmentation, Feature Extraction, and Classification by Ensemble Model,

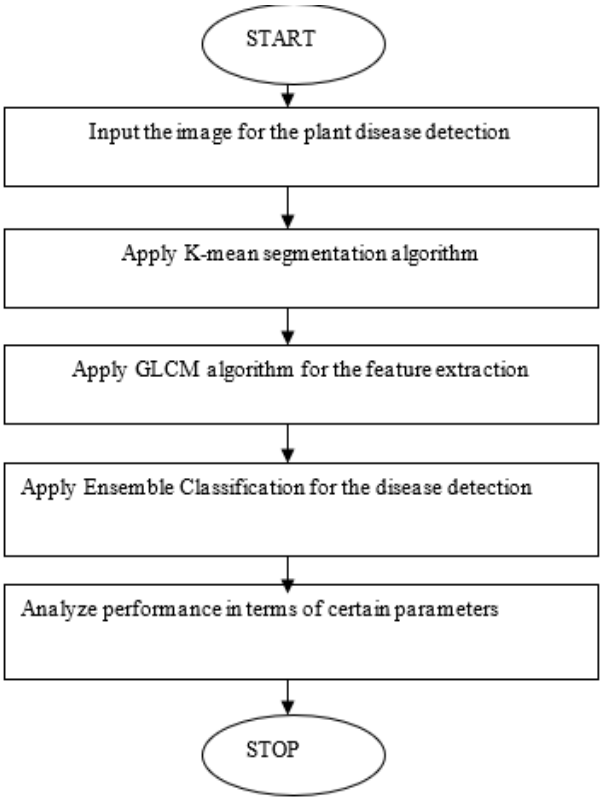

Fig. 3: Steps in the Proposed Work

3.1 Pre-processing: The primary stage aims to remove noise from the image. Here, the conversion of input leaf picture will be carried out into the gray scale. The noise will be removed from the image using the filtering method. In this phase of pre-processing image is filtered to remove noise from the image. The *Gaussian filter* is used to improve the noise and technique of *histogram equalizer* will improve contrast of the input image.

3.2 Segmentation: Segmentation is the approach which segments the image into certain segments. The region-based segmentation method called *k-mean clustering* is applied for this purpose here.

3.3 Feature Extraction: - The method of textural feature analysis called *GLCM* is applied in this research work. GLCM algorithm will extract *13 features* of the input images. These extracted features are: Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, Skewness, and IDM.

3.4 Classification: Here, ensemble classification method has been applied for the plant leaf disease detection. This ensemble classification method is obtained by the integration of two classification methods i.e. RF (Random Forest) and KNN method. RF technique has been applied for optimum feature selection, and then these features have been classified by KNN. These extracted features (obtained in step 3.3) have also been classified by SVM in order to assess the performance of the proposed approach.

4. RESULT AND DISCUSSION

This work is focused on the detection of plant disorder. The approach of ensemble classification is proposed to detect plant disorder. The ensemble classifier combines RF and KNN machine learning method. The performance of the proposed approach has also been evaluated with traditional support vector machine (SVM) on the basis of three measures such as accuracy, precision and recall.

***Accuracy*:** For a particular program, refers to the accuracy of the classification. It can be defined as the number of properly classified patterns to the total number of samples. The formula for accuracy measurement is mentioned below:

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

***Precision*:** The number of positive samples taking out from the total number of samples acknowledged positive by the classification model is called precision.

$$P = \frac{TP}{TP + FP}$$

***Recall*:** The number of true positive patterns taking out from the total number of positive declared patterns is called recall.

$$R = \frac{TP}{TP + FN}$$

Fig. 4 presents the *accuracy* comparison between the proposed ensemble classifier and SVM (existing traditional method). The ensemble classifier achieved the accuracy of 96% whereas SVM predicted the accuracy of 93%. Fig. 5 presents *precision* values achieved by SVM and the proposed ensemble classifier respectively. Also figure 6 presents *recall* values achieved by SVM and the proposed ensemble classifier respectively. Thus the experimental results revealed that the proposed approach outperformed SVM.

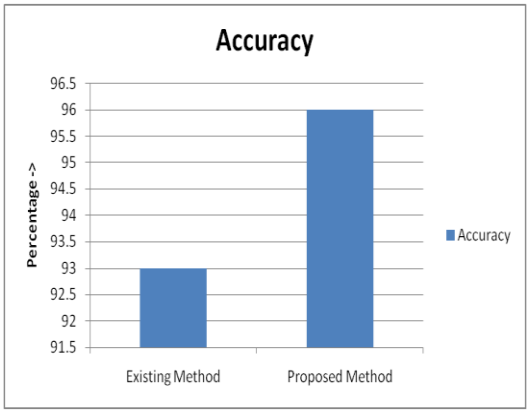

Fig. 4: Accuracy Comparison

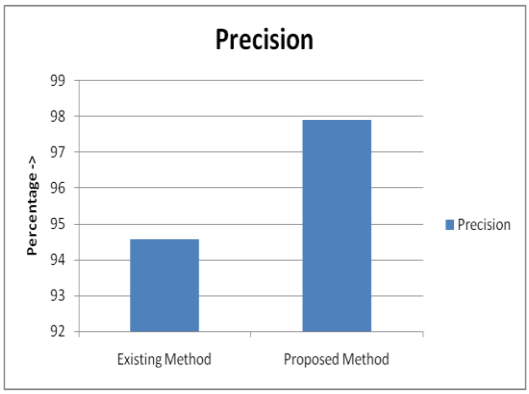

Fig. 5: Precision Comparison

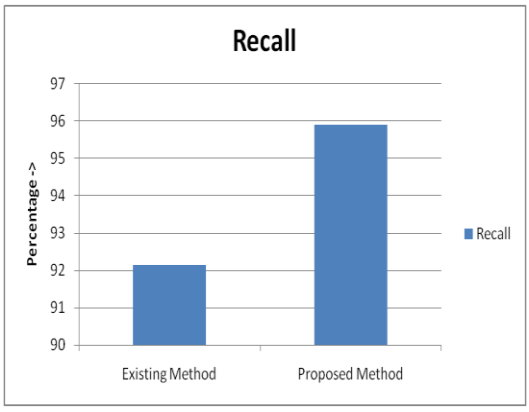

Fig. 6: Recall Comparison

5. CONCLUSION

The detection of plant disease is a major challenge for image processing and machine learning. The current work presents an ensemble approach for the detection of plant disease. This research work applies K-mean algorithm for the region-based segmentation. The GLCM is used for the textual feature extraction. The RF and KNN is used for the disease detection. The proposed approach has been found to be effective for disease detection with good accuracy.

References:

[1] Rekha Balodi, Sunaina Bisht, AbhijeetGhatak and K.H. Rao, "Plant disease diagnosis: technological advancements and challenges", Indian Phytopath. 70 (3): 275-281 (2017)
[2] MonishankerHalder, Ananya Sarkar, HabibullahBahar, "PLANT DISEASE DETECTION BY IMAGE PROCESSING: A LITERATURE REVIEW", 2019, SDRP Journal of Food Science & Technology
[3] Surender Kumar, Rupinder Kaur, "Plant Disease Detection using Image Processing: A Review", International Journal of Computer Applications (0975 – 8887) Volume 124 – No.16, August 2015
[4] Sabah Bashir, Navdeep Sharma, "Remote Area Plant Disease Detection Using Image Processing", IOSR Journal of Electronics and Communication Engineering (IOSRJECE), Volume 2, Issue 6 (Sep-Oct 2012), PP 31-34
[5] Rajneet Kaur, Miss. Manjeet Kaur, "A Brief Review on Plant Disease Detection using in Image Processing", International Journal of Computer Science and Mobile Computing, Vol. 6, Issue. 2, February 2017, pg.101 – 106
[6] Sushil R. Kamlapurkar, "Detection of Plant Leaf Disease Using Image Processing Approach", International Journal of Scientific and Research Publications, Volume 6, Issue 2, February 2016
[7] Vijai Singh, Varsha, A K Misra, "Detection of unhealthy region of plant leaves using image processing and genetic algorithm", International Conference on Advances in Computer Engineering and Applications, Year: 2015 | Conference Paper | Publisher: IEEE
[8] AbiramiDevaraj, KarunyaRathan, SarvepalliJaahnavi, K Indira, "Identification of Plant Disease using Image Processing Technique", International Conference on Communication and Signal Processing (ICCSP), Year: 2019 | Conference Paper | Publisher: IEEE
[9] R. Meena Prakash, G.P. Saraswathy, G. Ramalakshmi, K.H. Mangaleswari, T. Kaviya, "Detection of leaf diseases and classification using digital image processing", International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Year: 2017 | Conference Paper | Publisher: IEEE
[10] Sharath D M, Akhilesh, S Arun Kumar, Rohan M G, Prathap C, "Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight", International Conference on Communication and Signal Processing (ICCSP), Year: 2019 | Conference Paper | Publisher: IEEE
[11] Gurleen Kaur Sandhu, Rajbir Kaur, "Plant Disease Detection Techniques: A Review", International Conference on Automation, Computational and Technology

Management (ICACTM), Year: 2019 | Conference Paper | Publisher: IEEE
[12] Sachin D. Khirade, A.B. Patil, "Plant Disease Detection Using Image Processing", International Conference on Computing Communication Control and Automation, Year: 2015 | Conference Paper | Publisher: IEEE
[13] R Anand, S Veni, J Aravinth, "An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method", International Conference on Recent Trends in Information Technology (ICRTIT), Year: 2016 | Conference Paper | Publisher: IEEE
[14] Chaitali G. Dhaware, K. H. Wanjale, "A modern approach for plant leaf disease classification which depends on leaf image processing", International Conference on Computer Communication and Informatics (ICCCI), Year: 2017 | Conference Paper | Publisher: IEEE
[15] UmutBarişKorkut, ÖmerBerkeGöktürk, OktayYildiz, "Detection of plant diseases by machine learning", 26th Signal Processing and Communications Applications Conference (SIU), Year: 2018 | Conference Paper | Publisher: IEEE
[16] Pranjali B. Padol, Anjali A. Yadav, "SVM classifier based grape leaf disease detection", Conference on Advances in Signal Processing (CASP), Year: 2016 | Conference Paper | Publisher: IEEE
[17] RashmiPawar, AmbajiJadhav, "Pomogranite disease detection and classification", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Year: 2017 | Conference Paper | Publisher: IEEE