Plant Leaf Disease Detection Approach Based On Color and Texture, Using KNN AND SVM.

Dr. Pankaj Mohindru[1], Gagandeep Kaur[2], Dr. Pooja[3]
Assistant Professor, Department of Electronics, Punjabi University, Patiala, India[1]
M.Tech Scholar, Department, of Electronics, Punjabi university, Patiala, India[2]
Assistant Professor, Department of Electronics, Punjabi University, Patiala, India[3]

Abstract: The sensible and precise analysis of plant diseases plays a significant role in inhibiting the loss of yield and loss or reduced quantity of agricultural products. In current years, machine learning is extensively used in image processing. In this paper, we evaluated the performance results using powerful architectures of neural networks for plant disease detection. In this paper SVM and KNN is implemented which contains two datasets; one is training dataset and train dataset. Initially original image is captured and is used for further analysis. It gives us the color and background segmented pixel set, also separated the saturation and hue part of test image. This segmented image is further analyzed to detect diseased part of the leaf. This work will also provide percentage of area in which diseases is present and give us the detail of bacteria that is infecting the leaf. The evaluation results show that machine learning methods and KNN classification give better outcome than proposed algorithm.

Keywords: *Image Samples of plant leaf, Image Recognition, Artificial Neural Network, KNN, SVM.*

I. INTRODUCTION

Agriculture is a vital source of a country economy and it is the intermediate of living of almost nearly two third population of a world. The economy depends on the quality and quantity of agricultural products and the quality and quantity is subject to a good weather and other environmental factors. As varieties of agriculture products are produced and exported to different countries, thus it is necessary to harvest high quality products with an equitable yield. Plant disease is one of the most indicated sign for the degradation of products quality [1]. In recent days disease vindication would be considered to be an influential factor in the case of agriculture. As, the diseases of the plants are unavoidable, detection of plant disease are essential in the field of farming. The coolest way of finding a disease affected plant is done by examining its leaf condition. The prime diseases of plants are due to viral, fungus and bacterial diseases such as anthracnose, bacterial blight, canker, and leaf spot etc. Plant disease detection is a very thought-provoking topic for agriculture specialists which requires the adoption of scientific methods and longtime of observance. It is not an easy task to detect a plant disease by examining only its leaves because many diseases have similar symptoms. Thus, there is a scarcity of an automated system that can perform the operations of plant disease detection and gives an effective solution. To accomplish this target image processing techniques are extensively used [2]. Color image segmentation plays an important role in color image analysis. Image segmentation is the process of extracting an area of concern from an image. There are exists numerous color image segmentation techniques and KNN clustering is a widely used color image segmentation algorithm [3]. KNN clustering is a process of arranging the similar pixels into k number of clusters. The selection of different cluster gives different segmentation result of a picture and it becomes difficult to find the appropriate cluster number. Various segmentation and machine learning algorithms are lately offered for the recognition and classification of plant diseases from leaf images. These techniques have created a path to remove the problems but the challenged being faced is the performance of the results obtained. This paper approaches a method for plant disease detection and classification using k-nearest neighbor and SVM classifier.

Analysis of plant diseases and its indicators

For a healthy growth of plants, it is very important to know about the various infections on plants caused by different diseases**.** The use of visible-wavelength image is one of developing methods to detect and enumerate disease on plants. Apart from traditional visual valuation methods used to recognise and estimate disease intensity, several other methods based on various studies are also used, including laser-induced fluorescence, radar, microwave, thermography, nuclear magnetic resonance imaging and multi or hyper-spectral imagery [6]. Various forms of microscopy are also used to recognise and count various pathogens in infected plants. Awareness of these approaches is vital and RGB image feature pixel counting techniques is broadly used to agricultural science [9]. Image examination can be applied for the following purposes:

- To detect plant leaf, stem, and fruit diseases.
- To detect affected area by disease.
- To find the area of the affected plant.

II. SUMMARY OF THE TECHNQIUES TO BE USED

Global Image Descriptor Task (GIST) [13] - Global features describe the image as a whole to take a broad view of entire object. GIST features include contour representations, shape descriptors, and texture features. Shape Matrices, Invariant Moments, Histogram Oriented Gradients and Co-Histogram Oriented Gradients are some examples of GIST.

The Scale-Invariant feature transform (SIFT) [15] is a feature detection algorithm in computer vision to detect and describe local features in images. SIFT key points of objects are first extracted from a set of reference pictures and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors

Histogram of Oriented Gradient (HOG) [12] is used for object-detection. It is based on computing edge- gradients. HOG used the sliding window method for object detection. HOG divide the analyzed image into the smaller cells, called blocks. HOG can be used alongside SVM and KNN for feature detection.

CENTRIST- it is a tool for visual description and is very useful in areas where illumination changes, gamma variation etc. happened in comparison to GIST and SIFT [8]. CENTRIST is histogram of Census Transform values. Census Transform assigned 0 and 1 to the neighboring pixel. The decimal number corresponding to this sequence of eight neighbouring binary digits is computed and used as census transform value of Centre pixel.

III. PROPOSED METHODOLOGY

Scene classification system involves various steps like feature extraction, feature selection, feature vector generation, training and classification. The flow chart in figure 1 demonstrates all the phases system goes through. The data preparation phase involves the gathering the leaf image dataset from available sources. The actual design of the system starts with collecting data for testing. The collected images are then transformed into an appropriate format suitable for system design.

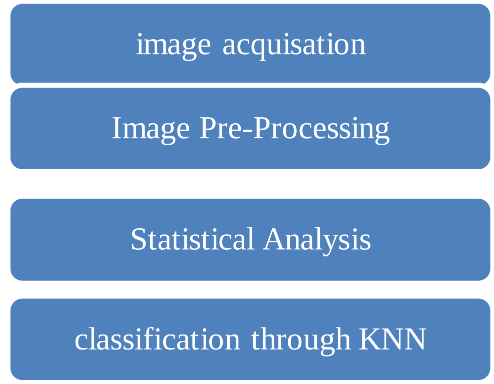


TABLE 1: The description of the leaf disease classification phases

(a) Bacterial disease (b) viral disease

READ INTO IMAGES INTO MATRIX

RESIZE IMAGES

CONVERT RGB IMAG INTO GREY SCALE

CREATE IMAGE FILTER

APPLY FILTER ON ALL IMAGES TO REMOVE BACKGROUND NOISE

PREPROCESSED IMAGES

Table 2 Preprocessing phase

Preprocessing phase

Preprocessing involves following steps:

1. Read image into matrix: **-** Reading the image into matrix is the very first step involved in image processing. The size of matrix is proportional to size of image. Each cell of matrix stores the corresponding pixel value of image.

2. Resize the image: **-** All images from training and testing dataset are normalized to same size.

3. Convert RGB images to gray-scale: **-** The RGB images are converted into gray scale images to make the model computationally efficient.

4. Create image filter:-The Gaussian filter is applied on the image in order to remove the Gaussian noise from the leaf disease test object after converting it into the gray scale.

A Gaussian filters smoothen an image by calculating weighted averages in a filter box. Gaussian filtering is used to blur images and remove noise and detail. In Gaussian filter following Gaussian function is used to calculate the weights:

$$g(i, j) = c \cdot e$$

The properties of Gaussian Filter are as following:

- To reduce the Gaussian noise, the filter for Gaussian smoothing is considered efficient.
- To reduce the blurring effect on the edges of the images based on the calculation histogram based weights.
- To remove the noise in all of the directions or dimensions in the image object.
- The degree of smoothing is controlled by σ (larger σ for more intensive smoothing).

5. Apply filter: **-** The filtered images become more robust to feature extraction and classification. The image is convolved with Gaussian filter.

(a) infected leaf

(b) Diseased part segmented from original image

Feature extraction algorithm-

- Load the colored test image as test object matrix.

- Convert the test image matrix to grey scale image matrix.
- Define the Gaussian Filter.
- Apply Gaussian Filter on grey scale image to produce the de-noised image.
- Calculate $\mathbf{G_{md}}$ into the front-ground estimation feature.
- Define the dilation object of adequate shape and size.
- Dilate the image $\mathbf{G_{md}}$ with respect to shape and size to produce the image object $\mathbf{A_I}$

Training:-

- Validate image size:- Each training image is resized to 128×128 dimensional matrix. Image resize is mandatory to keep each feature vector of same size.
- Create group vector:- Group vector here signifies a vector containing target values of each indoor class. Group vector contains five category labels, one of each class. Each sample from a specific class is assigned same category label to distinguish these from other class labels. Group vector is 1×125 dimensional vector with one category label for each of 125 training images.
- Create training set: - Training set is created to be fed to learning model. Training set is having dimensions 125×48000 for 125 training images and 1×48000 dimensional feature for each image.
- The one-vs.-all classifier has been used for this multiclass classification. There are two inputs to training module: 1) 125×48000 dimensional training set with each row corresponds to one observation and column corresponds to one feature. 2) 1×125 dimensional group vector, where each element specifies category label corresponding to each row of training set.

Learnt model: - KNN model have been learnt for each class where each model can distinguish a particular class images from other.

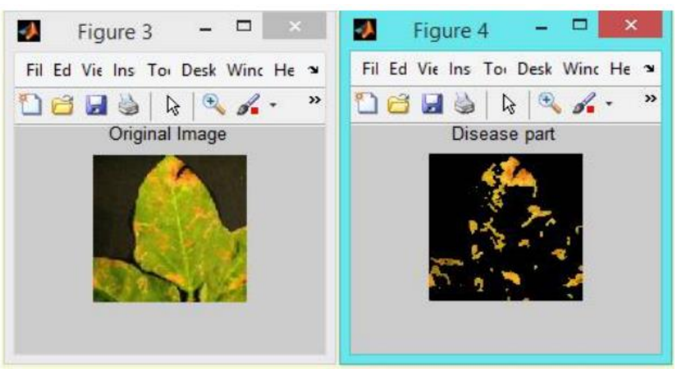


Fig Classification of result displaying in the massage box

Testing

- Extract features for test image: - The features are extracted from test image.
- Classify: - The test image is fed to learn model. The five learnt models each assign binary 1 or 0 to test image. The test mage is assigned a category label.

IV. DATA AND RESULTS

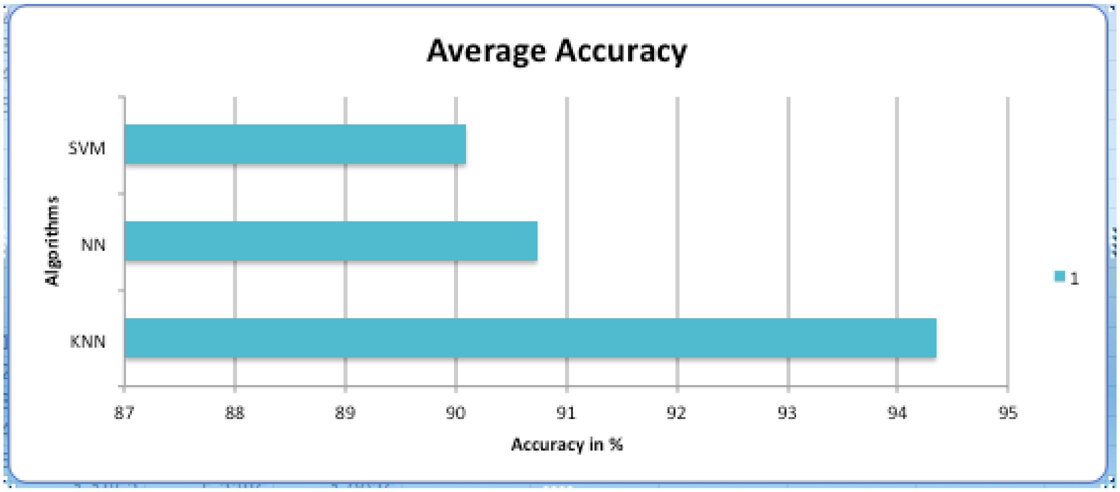
Database collection - Database of 100 images of leaves of infected and healthy plants are captured in jpeg format.

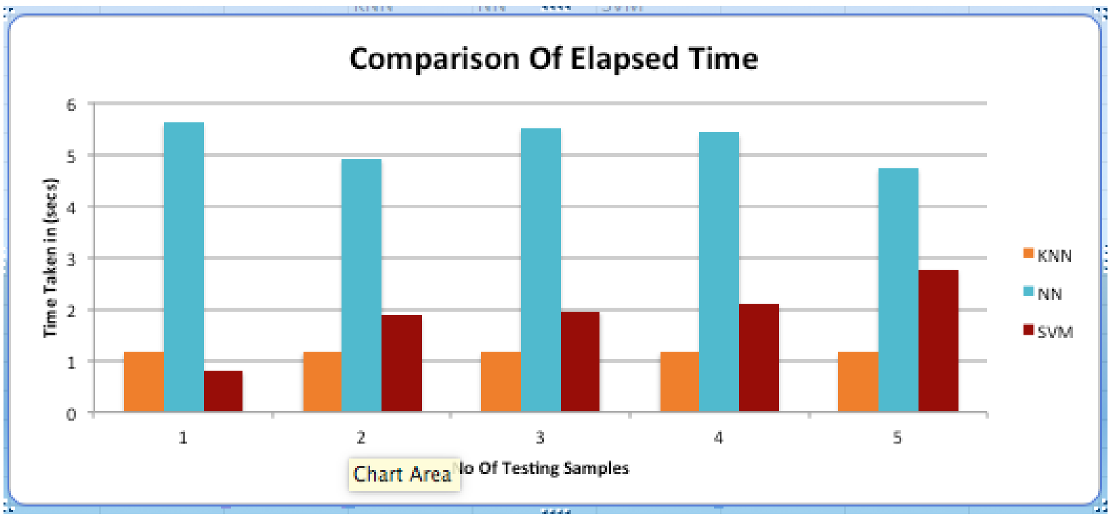
V. RESULT ANALYSIS

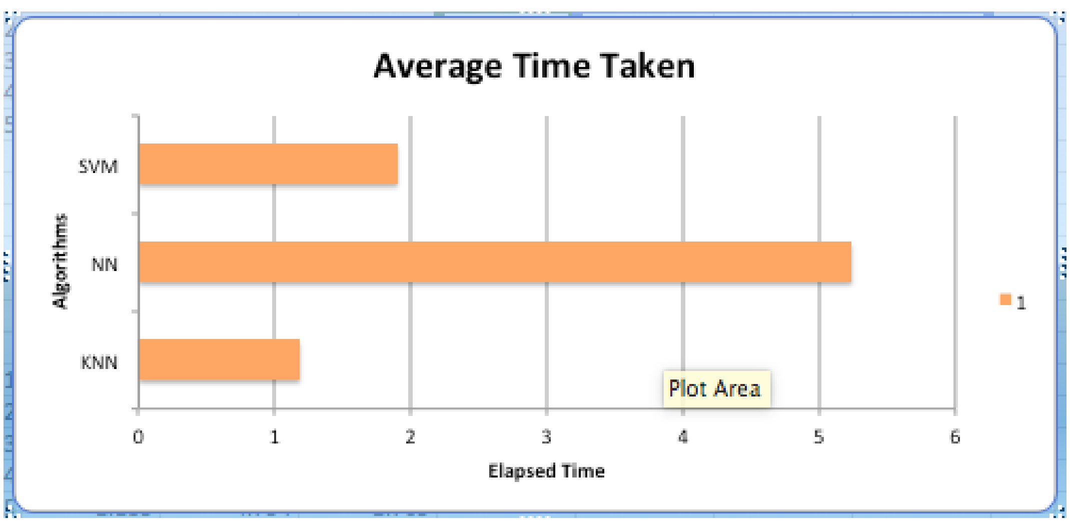
The results have been obtained from the various experiments conducted over the proposed model. The training dataset has been classified into two primary classes, which primarily defines the given image is proved and the decision logic is returned with the detected category type.

Sr. No.	KNN	NN	SVM
1	1.905	6.511	0.908
2	1.277	5.923	2.806
3	1.287	6.313	1.995
4	1.281	5.512	2.565
5	1.290	4.634	2.665

Sr. No.	KNN	NN	SVM
1	94.65	91.82	89.80
2	94.93	91.23	89.94
3	94.95	90.28	90.23
4	94.96	90.42	90.54
5	94.36	89.91	89.95


Average Accuracy
SVM
NN
KNN
Algorithms
87
88
89
90
91
92
93
94
95
Accuracy in %
1


Comparison Of Elapsed Time
Time Taken in (secs)
0
1
2
3
4
5
6
1
2
3
4
5
KNN
NN
SVM
Chart Area
No Of Testing Samples


Average Time Taken
SVM
NN
KNN
Algorithms
0
1
2
3
4
5
6
Elapsed Time
1
Plot Area

Above graph represents the average time taken by each system in the testing phase

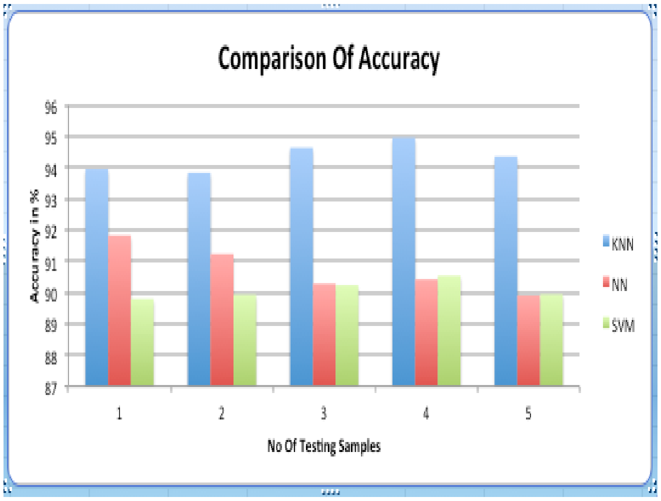


Fig 4 Total Time Taken by all Three Classifiers

Above chart is the comparison chart for elapsed time for processing and categorize the disease found in the testing image. Results show that KNN has taken less time for computation. It means KNN is faster in comparison of NN and SVM.

VI. CONCLUSION

In this paper, we will go to examine the different favorable circumstances and disservice of the plant infections forecast procedures and going to propose a novel approach for the recognition calculation. The images of different leaves are gained utilizing an advanced camera. At that point picture handling systems are connected to the gained pictures to separate helpful elements that are important for assist examination. From that point onward, a few logical systems are utilized to order the pictures as per the particular issue close by.

The proposal reviews and summarizes some techniques that have been used for plant disease detection. The two major classification techniques of plant diseases .The spectroscopic and imaging techniques include fluorescence spectroscopy, visible-IR spectroscopy, fluorescence imaging, and hyper spectral imaging. KNN algorithm is effective classifier has been used here to minimize the computation cost. Also used by many researchers used this classifier on many datasets. KNN classifier obtains highest result as compared to SVM. The comparison will be done on accuracy and detection time based parameters and will show that KNN is better than existing SVM. A novel approach for disease prediction of plants based on classification technique is proposed.

VII. FUTURE WORK

In the future the proposed model can be improved by using the hybrid low level feature extracted along with the efficient color illumination to find the dual-mode attacks over the images to determine the plant disease in the image data. Also the swarm intelligent algorithm can be utilized for the plant disease recognition in the digital image dataset. Further speed could be focused.

VIII. REFERENCES

[1]. Garima Tripathi and Jagruti, “An image processing and Neural network based approach for detection and classification of plant leaf diseases,” Volume 6, Issue 4, April (2015), pp. 14-20

[2]. P.Revathi, M.HemaLatha, “Classification Of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques,” ISBN, 2012, 169-173, IEEE

[3]. Santanu Phadikar& Jaya SiL, “Rice Disease Identification Using Pattern Recognition Techniques, Proceedings of 11th International Conference on Computer and Information Technology, 25-27.

[4]. Geng Ying, Li Miao,Yuan Yuan & Hu Zelin, “A Study on the Method of Image Pre-Processing for Recognition of Crop Diseases,” International Conference on Advanced Computer Control, 2008 IEEE.

[5]. Ajay A. Gurjar, Viraj A. Gulhane, “Disease Detection on Cotton Leaves by Eigen feature Regularization and Extraction Technique”, International Journal of Electronics, Communication & Soft Computing Science and Engineering (IJECSCSE) Volume 1, Issue 1.

[6]. H. Al-Hiary, S. Bani-Ah Mad, M. Reyalat, M. Braik and Z. A Lrahamneh, “Fast and Accurate Detection and Classification of Plant Diseases,” IJCA, 2011, 17(1), 31-38, IEEE-2010.

[7]. Tejal Deshpande, Sharmila Sengupta, and K.S.Raghuvanshi, “Grading & Identification of Disease in Pomegranate Leaf and Fruit,” IJCSIT, vol. 5 (3), pp 4638-4645, 2014.

[8]. P.Revathi and M.Hemalatha, “Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques,” IEEE International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET), Tiruchirappalli, pp 169-173, 2012. 59

[9]. Ms. Kiran R. Gavhale, Prof. Ujwalla Gawande, and Mr. Kamal O. Hajari, “Unhealthy Region of Citrus Leaf Detection using Image Processing Techniques,” IEEE International Conference on Convergence of Technology (I2CT), Pune , pp 1-6, 2014

[10]. Monika Jhuria, Ashwani Kumar and Rushikesh Borse, “Image processing for smart farming detection of disease and fruit grading,” IEEE 2nd International Conference on Image Information Processing (ICIIP), Shimla, pp 521-526, 2013

[11]. Evy Kamilah Ratnasari, Mustika Mentari, Ratih Kartika Dewi and R. V. Hari Ginardi, “Sugarcane Leaf Disease Detection and Severity Estimation Based On Segmented Spots Image,” IEEE International Conference on Information, Communication Technology and System (ICTS), Surabaya, pp 93-98, 2014

[12]. Ratih Kartika Dewi and R. V. Hari Ginardi, “Features Extraction for Identification of Sugarcane Rust Disease,” IEEE

International Conference on Information, Communication Technology and System (ICTS), Surabaya, pp 99-104, 2014
[13]. Vani Ashok, and Dr. D.S. Vinod, "Automatic Quality Evaluation of Fruits Using Probabilistic Neural Network Approach," IEEE International Conference on Contemporary Computing and Informatics (IC3I), Mysore, pp 308-311, 2014
[14]. Dheeb Al Bashish, Malik Braik, and Sulieman Bani Ahmad, "Feature Extraction for Identification of Sugarcane Rust Disease," IEEE International Conference on Signal and Image Processing (ICSIP), Chennai, pp 113-118, 2010.
[15]. Aakanksha Rastogi, Ritika Arora and Shanu Sharma, "Leaf Disease Detection and Grading using Computer Vision Technology & Fuzzy Logic," IEEE 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), Noida, pp 500-505, 2015.
[16]. Juneja, Mayank, Andrea Vedaldi, C. V. Jawahar, and Andrew Zisserman. & quot; Blocks that shout: Distinctive parts for scene classification." In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pp. 923-930. IEEE, 2013.
[17]. Monadjemi, Amir, B. T. Thomas, and Majid Mirmehdi. Experiments on high resolution images towards outdoor scene classification. Technical report, University of Bristol, Department of Computer Science, 2002.
[18]. Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley & Sons,, 1999.
[19]. Fitzpatrick, Paul. "Indoor/outdoor scene classification project." Pattern Recognition and Analysis.
[20]. Quattoni, Ariadna, and Antonio Torralba. "Recognizing indoor scenes." In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 413-420. IEEE, 2009.
[21]. Li, Li-Jia, Hao Su, Yongwhan Lim, and Li Fei-Fei. & quot ; Objects as attributes for scene classification." In Trends and Topics in Computer Vision, pp. 57-69. Springer Berlin Heidelberg, 2012.
[22]. Antanas, Laura, Marco Hoffmann, Paolo Frasconi, Tinne Tuytelaars, and Luc De Raedt. & quot: A relational kernel-based approach to scene classification. & quot; In Applications of Computer Visio (WACV), 2013 IEEE Workshop on, pp. 133-139. IEEE 2013
[23]. Mesnil, Grégoire, Salah Rifai, Antoine Bordes, Xavier Glorot, Yoshua Bengio, and Pascal Vincent. " Unsupervised and Transfer Learning under Uncertainty-From Object Detections to Scene Categorization." In ICPRAM, pp. 345-354. 2013.
[24]. Zhang, Lei, Xiantong Zhen, and Ling Shao. "Learning object-to-class kernels for scene classification & quot; Image Processing, IEEE Transactions on 23, no. 8 (2014): 3241-3253.
[25]. Li, Li-Jia, Hao Su, Li Fei-Fei, and Eric P. Xing. & quot; Object bank: A high-level image representation for scene classification & semantic feature sparsification. & quot; In Advances in neural information processing systems, pp. 1378-1386. 2010.
[26]. Alberti, Marina, John Folkesson, and Patric Jens felt. & quot; Relational approaches for joint object classification and scene similarity measurement in indoor environments. & quot; In AAAI 2014 Spring Symposia: Qualitative Representations for Robots. 2014.
[27]. Russakovsky, Olga, Yuanqing Lin, Kai Yu, and Li Fei-Fei. & quot; Object- centric spatial pooling for image classification." In Computer Vision–ECCV 2012, pp. 1-15. Springer Berlin Heidelberg, 2012.
[28]. Espinace, Pablo, Thomas Kollar, Nicholas Roy, and Alvaro Soto. &quot ;Indoor scene recognition by a mobile robot through adaptive object detection." Robotics and Autonomous Systems 61, no. 9 (2013): 932-947.
[29]. C. Fredembach, M. Schroder, S. Susstrunk, Eigenregions for image classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (12) (2004) 1645–1649.
[30]. A. Mojsilovic, J. Gomes, B. Rogowitz, Isee: perceptual features for image library navigation, in: SPIE Human Vision and Electronic Imaging, 2002.

Dr. Pankaj Mohindru, received the Master of Engineering Degree with distinction in Electronics from Thapar University, Patiala in the year 2005. He is awarded Ph.D. by Punjabi University, Patiala in 2011. Presently, he is working as Faculty in Electronics and Communication Engineering Department, Punjabi University, Patiala. His research area include Fuzzy Logic, Measurement Science & Techniques & Research Methodology etc. He has 35 research papers in reputed International Journals.

Ms Gagan is presently pursuing her M.Tech. in Electronics and Communication Engineering in the Department of Electronics and Communication Engineering at the Punjabi University, Patiala, Punjab, India. She obtained her B.Tech in Electronics and Communication Engineering from Patiala Institute of Engineering and Technology (PIET) affiliated to PTU. Her areas of interest include Fuzzy Logic and Neural Netweorks.

Dr. Pooja received B.E. Honors Degree in Electronics & Communication in 2001 from Deen bandhu Chhotu Ram University of Science and Technology, Murthal and M.E. Degree in 2003 from Thapar University, Patiala. Dr Pooja is awarded Ph.D. in E.C.E. from Thapar University in 2014. Presently, she is working as Assistant Professor in Electronics and Communication Engineering Department, Punjabi University, Patiala. Her research area includes Signal processing and Digital communication. She has 20 research papers in reputed national/international journals. Five of them are published in science citation index journals.