International Journal of Pure and Applied Mathematics
Volume 117 No. 15 2017, 699-710
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

PADDYLEAF DISEASEDETECTION USING SVM WITHRBFNCLASSIFIER

T. GAYATHRI DEVI[1], P. NEELAMEGAM[2]
[1]Department of ECE, School of EEE, SRC, SASTRA University, Kumbakonam
[2]Department of E&I, School of EEE, SASTRA University, Thanjavur

Abstract

Now-a-days disease damage is one of the most important problems in the agriculture sector especially in the paddy leaf because it gradually reduces the yield also minimize the health condition of the rice. The problem present in the agriculture sector is minimized by applying the various image processing and machine learning techniques but the certain situation, the elimination of the disease is still issues. So, in the paper introduces the novel automated rice leaf disease identification and detection system using the improved support vector machine with the radial basis neural networks. Initially the captured paddy images are transformed into the gray scale image and the noise present in the image is eliminated with the help of the image clipping, cropping and smoothing process. Afterwards, the image enhancement need to perform by applying the histogram equalization method and the particular affected region is segmented with the help of the Otsu' thresholding with k-means clustering process. From the segmented region, different features are extracted using scale invariant method and the features are classified with the help of the support vector machine trained radial basis neural network. The outcome of the key parameters such as, error rate, sensitivity, specificity and accuracy will evaluate the efficiency of the system.

Keywords:

Disease damage, paddy leaf, gray scale image, image clipping, cropping,image enhancement, Otsu' thresholding, k-means clustering, scale invariant, support vector machine trained radial basis neural network.

1. Introduction

India is a cultured country and the 70% of the people's basic day to day life only depends on the agriculture. Based on the importance of the agriculture, farmers are chosen their paddies, crops and the related pesticide to improve the growth of the plant in the limited period. The abnormal usage of the chemical pesticides, mechanical pesticides, nutritional problems and cultural problems increases the damage in the leaves. The damages may occur in terms of the fungi, viruses, viroid, viruses, bacteria, phytoplasma, insects, nematodes and mites. The factors of the damage is reduces the both quality and quantity of the agriculture product gradually [1]. Then the sample affected paddy leaf shown in the figure 1. Various diseases are occurred to the paddies among that paddy blast and brown spot diseases are more severe diseases because it not having the clear symptoms initially. Magnaportheoryzae, the fungus influences the blast in paddiesin which the paddies are appeared initially in the green colour. Afterwards, the grey-green colour is changed into the dark green spots [2]. The lesions on the leaves appear in any one of the following outlook:

1. Elliptical or spindle shape with whitish to grey centres surrounded by red to brownish

2. Diamond shape with wide centre and pointed towards end spot.

Leaf blast lesions are elongated and pointed at each end of the spot. The best disease management strategies are diseased-free seeds, resistant cultivars and proper plant spacing, transplanting [than broad casting] and split applications of nitrogenous fertilizer. When plants are wet, it is better to avoid farm activities. The Helminthosporiumoryzae [3] fungus causes the small, circular, and dark brown to purple-brown Brown spot.The brown spot diseases change the colour of paddies from lightbrown into the grey centre surrounded by the reddish brown margin. This will create the serious effects when compared to the other paddy disease.

Figure 1: Sample Diseases Affected Paddy Leaf

Based on the severity of the paddy disease, various researches uses the automatic detection system for recognizing the affected paddies using the different machine learning techniques like median filter, Weiner filter, non-local noise filter, Gaussian filter, clustering methods, edge detection, feature extraction, k-nearest neighbour and neural network methods are used. This method effectively identified the affected paddies but sometimes, the symptom fewer paddies are difficult to identify. The difficulties lead to reduce the entire system performance like error rate, accuracy and so on [4]. So, the proposed system is automatic paddy disease recognition system using the SVM trained radial basis neural network. Before classifying the features, the affected region is efficiently segments with the help of the Otsu with k-means clustering process. The optimized region segmentation process successfully determines the diseases present in the paddies. Finally the performance of the system is evaluated with the help of the Rice knowledge bank dataset [5]. The collected dataset images are preprocessed and the extracted features are classified which ensures whether the paddies are affected by disease or not. Then that the obtained results are compared with the various methods like KNN, MLP, and SVM.

The later part of this paperdiscussesthe related reviews about paddy leaf diseases. Proposed automatic paddy disease identification system, the efficiency of the suggested system by the author and conclusion of the work.

2. Related Works

In this section discusses about the various research methodologies about the paddy disease identification process. NunikNovianaKurniawati et al., [6] has deduced an image processing technique for automatic, detection and classification of various paddy diseases.The author proposes in this paper,the disease detection part that uses Haar-like feature and AdaBoost (Adaptive Boosting) classifier to locate the disease affected portion of the paddy plant. Disease recognition part uses SIFT. At some position and scale the pixels will be aggregated inside a basic rectangular Haar like component .This changed list of capabilities is known as 2-rectangle highlight. The value indicates certain characteristics of a particular area of the image.

Meunkaewjinda et al., [7] analyzes the Hybrid Intelligent Systemto obtain more value added and high quality agricultural export products of Thailand. Multiple non-natural intelligent techniques are used for an automatic plant infection analysis using Leaf feature inspection mainly used for infection of disease. The system can analyse or diagnose plant leaf disease without further maintenance and expertise if the system is trained.. Grape leaf colour segmentation, Grape leaf disease segmentation has been used for analysis the disease in the system.

Arivazhagan et al., [8] has developed four principle steps are first a colour transformation structure for the input RGB image is generated, and then the green pixels are marked and eliminated using specific threshold value followed by segmentation process, evaluating the texture features using colour co-occurrence technique for the useful segments, finally the extracted features are proceed through the classifier. A related and supervised learning method is better known as SVM, which is utilized for the process of classification and regression. The detection accuracy is enhanced by SVM classifier. The two class problem is then expanded to multi class problem where the detected leave diseases are then organized into various groups. By this technique, the plant diseases can be detected at initial stage itself and the pest control techniques can be used to resolve pest problems while minimizing risks to people and the environment.Amit Kumar Singh et al., [9] analyzing the paddy disease in the rice plant using the support vector machine. The author analyzes the leaf which is collected from the international rice research institute. The collected leaf images are segmented with the help of the k-means clustering process. From the segmented image various features are extracted which are classified by applying the support vector machine approach. The system ensures 82% accuracy which designates the efficiency of the system. According to the above discussions the proposed paddy disease identification system is implemented as follows.

3. Proposed Methodology

In this section, we have discussed the proposed paddy leaf disease detection methodology because the paddy diseases are mostly affects the agriculture growth in terms of both quantity and quality [10]. The proposed automatic system improves the disease recognition rate because it uses the efficient image processing and machine learning stages like image collection, noise removal, image segmentation, feature extraction and feature classification. Based on the above stages, the proposed system working structure [11] is shown in the figure 2.

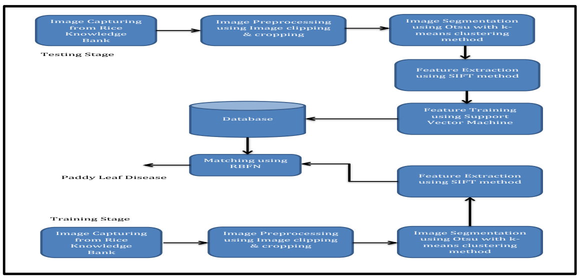

Figure 2: Proposed System Architecture

The above figure 2 clearly shows that the designed system architecturehas two different stages . one is training and the other is testing stage. In the training stage the images are collected from the Rice knowledge bank dataset and the noise has been removed by applying the image clipping, cropping methods. Then the particular affected region is segmented with the help of the Otsu thresholding with the k-means algorithm[18]. Afterwards the SIFT features are extracted which are trained with the help of the support vector machine which is stored in the database. In the testing stages all the steps are performed like as the training stage expect the training process and matching is performed with the help of the radial basis neural networks. Then the detail discussion about the proposed methodology is explained as follows.

Training Stage

3.1 Image Preprocessing

The first stage of the paddy leaf disease identification is image pre-processing which is done with the help of the cropping, clipping and other process. Before processing the image need to be converting into the grayscale image because it provides the better results when compared to the colour image processing. The grayscale image contains the black, white and in between gray colours [12]. In that grey scale image, the white pixel consist of (255,255,255), black pixel has (0,0,0) and the middle value grey pixel is (127,127,127). According the to the pixel value, the gray scale value is estimated using the average weighted value of the red, green and blue value. The estimation process of the greyscale value is defined as follows.

$$GS = 0.2989 * Intensity(r) + 0.58701 * Intensity(g) + 0.1140 * Intensity\ (b) \quad (1)$$

According to the above equation (1), the image colour has been transmitted and the clipping and cropping process is performed which is further enhanced with the help of the histogram equalization process. The histogram process improves the contrast in the agriculture image and limits the amplification by using the clipping histogram [13]. The clipping value is determined with the help of the number of region wants to be enhanced using the contrast. Improving the contrast of the image is determined using the rayeigh distribution function which is defined as follows.

$$Rayleighg = g_{min} + \left[2(a^2)\ln\left(\frac{1}{1-p(x)}\right)\right]^{0.5} \quad (2)$$

Where, g_{min} is the minimum pixel value and *p(x)* is the cumulative probability distribution. *a* is the clip limit value. According to the above process the image has been enhanced and fed into the next image segmentation process.

3.2 Image Segmentation

The next stage is image segmentationwhich is the process of partitions an image. It's the process of segment the similar attributes into image which is done with the help of the Otsu with k-means clustering approach. The Otsu method [14] develops the two fold image from the gray level image according to the location of every pixel which is below to the threshold value and all the pixels above that threshold to one. Based on the threshold value the pixels are divided into two clusters that is one is belongs to the normal and other one is abnormal region. Then the mean valueof the particular pixels are calculated according to the mean value the clusters are formed and the distance between the features are reduced with the help of the k-means algorithm and the related objective function. The objective function is defined as follows.

$$argmin \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (3)$$

Where, x is the each pixel present in the image, $\mu_i ismeanvalueoftheparticularpixels.$

According to the above process, the clusters are dividing the paddy leaf regions into different regions. Then the sample segmented otsu based k-means clustering image is shown in the figure 3.

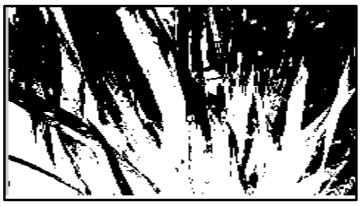

Figure 3: Thresholding based clustered Image

3.3 Feature Extraction

The next stage is feature extraction .the scale invariant feature transform is used to derive important features from the segmented region [15]. The method retrieves the feature according to the relative position because it does not change from one image to another image. It uses the key point detection, key point location, orientation assignment and key point descriptors techniques for feature extraction. Initially the key point is determined using the Gaussian filter. Then the maximum and minimum value of the image is estimated from the segmented image as follows,

$$D(x, y, \sigma) = L(x, y, K_i\sigma) - L(x, y, K_j\sigma) \quad (4)$$

Where $D(x, y, \sigma)$the difference of the Gaussian image is, $L(x, y, K\sigma)$ is the convolution value of the image, $I(x, y)$ is the Gaussian blur value,

$$L(x, y, K\sigma) = G(x, y, k\sigma) * I(x, y) \quad (5)$$

Then the key point is located to evaluate the exact feature of the image. Based on the key point position, using the Taylor series the location of the key point scale is calculated, which is obtained by

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (6)$$

Then the orientation has been assigned as follows, which is used to identify the direction of the particular key point, measured by the magnitude and orientation estimation.

$$m(x, y) = \sqrt{\left(L(x+1, y) - L(x-1, y)\right)^2 + \left(L(x, y+1) - L(x, y-1)\right)^2} \quad (7)$$

$$\theta(x,y) = atan2\big(L(x,y+1) - L(x,y-1)\big), \big(L(x+1,y) - L(x-1,y)\big) \quad (8)$$

Where, $m(x,y) = magnitudeofthekeyimage$, $\theta(x,y) = orientationthekeypointimage$

Finally the key point descriptors are extracted by analysing the key point detector and the orientation assignment process. In which the image has been divided into the 4x4 histogram orientation and each orientation has the 16x16 region of the key points which has 8 bins and 28 elements. Those elements are normalized by using the threshold value 0.2. The key point which lies within the threshold value is considered as the paddy leafs features. In addition, the standard deviation, mean, texture, shape of the leafs are calculated which are fed into the training stage to effectively detect the disease present in the image.

3.4 Feature Training

The extracted features are fed into the feature training process which is done with the help of the support vector machine. The trained features are provide the efficient results while analysing the different number of features in different position. The SVM [16] is one of the supervised learning methods that minimize the matching error rate also maximize the geometric margin. Considered the database has D feature set in which those features are contains the output value as {1,-1}. The feature which is having the output value -1 that is represented as the disease related features and the whose value having 1 is belongs to the normal feature. If the dataset D has the n samples which is represented as follows.,

$$D = \{(x_i, y_i) | x_i \in R^p, y_i \in \{-1,1\}\} \quad (9)$$

During the training process the hyperplane is decided by using the following equation 10.

$$w.x - b = 0 \quad (10)$$

Where x is a feature in the database, w is the normal vector of the hyperplane and the b is offset of the hyperplane. Then the maximum geometric margin is reduced by using the above equation 10 which classifies the features into the matching and unmatching condition and the related outputs are stored in the database for template matching.

Testing Stage

In the testing stage, the sample paddy plants leafs are captured which are preprocessed and the affected regions are segmented according to the predefined methods in the training stage. Then the segmented regions are fed into the feature extraction process for extracting the different type of feature. The extracted features are matched with the training data set for detecting the disease present in the paddy leaf. Then the matching process is explained as follows.

3.5 Disease Classification

The final stage of the paddy leaf disease identification process which is done with the help of the radial basis function [17] based neural network that is one of the artificial neural network which consumes the non-linear input values and produces the linear output. The three layers of the network are Input layer, hidden layer and output layer. Itis the fast learning method while recognizing the paddy disease. Initially the input layer consumes the input (mean, median, variance, standard deviation etc) from the extraction steps which are the nonlinear input values. The consume inputs are transmitted to the hidden layer which utilize the rad bas or radial basis activation function that helps to analyse the output of the particular neurons. The hidden layers are selected depending on the number of features if the input layer has small number of neurons, then minimum two hidden layers are chosen for increasing the neural network processing else the hidden layers are chosen less than the number of neurons. Then the radbas activation function is represented as follows.

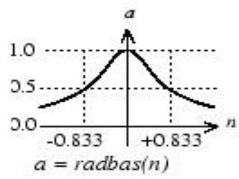

Figure 4: Representation of Radbas Activation Function

Every layer present in the network has two parameters like centers and spread factors. The system decides the centre value according to the k-means cluster which is used for disease diagnosis process. The k-means cluster is performed by using the equation 11. This is continuously updated for minimizing the error rate. Then the obtained result value is 1 then it is treated as the normal feature else it considered as the abnormal feature such as disease affected plant. Then the efficiency of the system is analysed using the experimental results and discussions which are explained as follow.

4. Performance Analysis

In this section discusses about the efficiency of the system using the rice knowledge database.

The total number of images trained in the data set paddy at different class.Total number of disease affected images (3 classes) = 120 images Total number of training samples = 90 images (each class contains 30 images) Total number of testing samples = 30 images (each class contains 10 images)The Training Phase, SIFT Oriented Gradient Features Are Extracted To Identify Two Classes Of Diseases Brown Spot And Leaf Blast Disease. Total disease affected images (2 classes)=90 images .Total number of trained images for 2 classes. For testing of images we collected the number of images for analysis the diseases.Then the performance of the proposed support vector training based radial basis function is analysed by the factors called, mean square error, and sensitivity, specificity and classification accuracy. The effective training process reduces the entire system error rate. Then the obtained mean square error value is shown in the following table 1.

Table 1: Mean Square Error of Different Classification technique

Classification Technique	Mean Square Error Value
BPN	0.89
RBFN	0.789
GRNN	0.934
SVM+RBFN	0.345

The above Table 1 clearly shows that the proposed Support vector trained with radial basis function neural network classifier has the minimum mean square error. Minimum Means Square Error Value is as shown in the Figure 5.

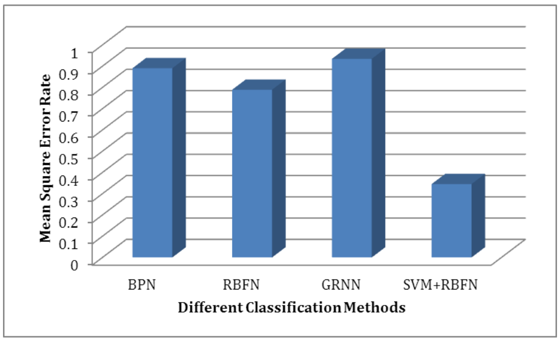

Figure 5: Mean Square Error of Different Classification technique

Then the proposed SVM with RBFN classifier classifies the paddy leaf disease from the captured paddy image with high sensitivity and specificity rate. The sensitivity and specificity value is calculated by using the following equation (11) and (12).

$$Sensitivity=TP/((TP+FN)) \quad (11)$$

$$Specificity=TN/((TN+FP)) \quad (12)$$

Where, TP = True Positive, TN = True Negative

FP = False Positive, FN = False Negative.

The Sensitivity and Specificity values of our system is compared to the several classification methods such as BPN, RBFN, and GRNN and shown in the figure 6.

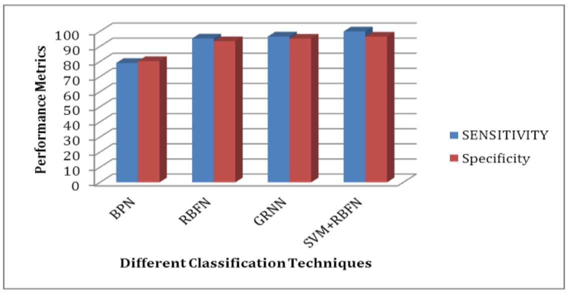

Figure 6: Sensitivity and Specificity of Different Classification technique

From the figure, it is easy to identify that the proposed system produces the best classification result which is shown in by using the sensitivity and specificity. Table 2 explains the classification accuracy of the suggested system.

Table 2: Classification Accuracy for Different Classification Techniques

S.No	**Classification Techniques**	**Classification Accuracy (%)**
1	BPN	85
2	RBFN	82.2
3	GRNN	85.6
4	SVM+RBFN	98.3

Thus the proposed system classifies the paddy disease from thecaptured images into abnormal and normal by using the SVM with RBFNwhich enhance the training also performance of the system. The above table shows the analysed different disease efficienc..

Table 3: Different Disease Accuracy Rate

Classifier	**Average Accuracy**	**Blast**	**Brown spot**
SVM	95.5	30	30
KNN	92.2	30	30

Thus the above discussions prove that the proposed system produces the highest accuracy compared to other existing recognition methods.

5.Conclusion

The farmers have to look after the crops from their early stages to identify the diseases. If the diseases are diagnosed well in early stages, it will be far better to treat them. Otherwise leads to heavy loss to the farmer and to the agricultural society. The machine vision system easily identifies the symptoms affected by paddy disease, through which farmers may have the properevaluation of crops in early stages. Sample images shown above exhibit the visual symptoms of a disease. K-means segmentation method is used to identify and segment the infected regions.SVM and RBFN classifiers uses input as colour and texture features extracted from each segmented region. The performance of SVM with RBFN classifier found to be better than other classifier for the proposed method. This work involves both image processing and pattern recognition techniques have significance to the real world categorization of crop disease. This workshould be extended to classify disease symptoms affected on fruits, vegetables, commercial crops etc.

6. References

[1] Al Bashish D,Braik M,Bani-Ahmad S (2010) "A Framework for Detection and Classification of Plant Leaf and Stem Diseases."In:2010 international conference on signal and image processing, IEEE, Chennai, pp 113–119

[2].DongPixia,WangXiangdong(2012)“Recognition of Greenhouse Cucumber Disease Based on Image Processing Technology” Open Journal of Applied Sciences, 2013, 3, 27-31

[3].Huang KY (2007) “Application of Artificial Neural Network for Detecting Phalaenopsis Seedling Diseases using Color and Texture Features.” Comput Electron Agric 57:3–11

[4]. Wang H, Li G, Ma Z, Li X (2012) “Application of Neural Networks to Image Recognition of Plant Diseases.”In: Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI). IEEE, Yantai, pp 2159–2164

[5]. H.A1-Hiary, S.Bani-Ahmad, M.Reyalat, M,Braik and Z. ALRahamnesh ,"Fast and Accurate Detection and Classification of Plant Diseases” ,IJCA, Vol: 17-No.1 ,March 2011.Pg: No: 31-37

[6]Nunik Noviana Kurniawati, Siti Norul Huda SheikhAbdullah, Salwani Abdullah, Saad Abdullah, ”Texture Analysis For Diagnosing Paddy Disease”, International Conference On Electrical Engineering And Informatics, 2009.

[7]]A.Meunkaewjinda, P.Kumsawat, K.Attakitmongcol &A.Srikaew[2008] Grape leaf disease detection from color imagery system using hybrid intelligent system, proceedings of ECTICON,2008,IEEE,PP-513-516

[8] S.Arivazhagan, R. Newlin Shebiah, S.Ananthi, S.Vishnu Varthini. (2013) “Detection of Unhealthy Region of Plant Leaves and Classification of Plant Leaf Diseases using Texture Features.” Agric Eng Int: CIGR Journal, 15(1): 211－217.

[9] Amit Kumar Singh, Rubiya, Senthil Raja, “Classification Of Rice Disease UsingDigital Image Processing And SvmClassifier”, International Journal of Electrical and Electronics Engineers, Volume 07, Issue 01, Jan-June 2015.

[10]John William Orillo, Jennifer Dela Cruz, Leobelle Agapito,Paul Jensen Satimbre Ira Valenzuela, “Identification OfDiseases In Rice Plant (Oryza Sativa) Using BackPropagation Artificial Neural Network”, 7th IEEEInternational Conference, 2013.

[11] Santanu Phadikar and Jaya Sil, „Rice Disease IdentificationUsing Pattern Recognition Techniques‟, Proceedings of 11thInternational Conference on Computer and InformationTechnology (ICCIT), 25-27 December 2008.

[12] Rong Zhou, Shun‟ichi Kaneko, Fumio Tanaka, MiyukiKayamori, Motoshige Shimizu, “Early Detection AndContinuous Quantization Of Plant Disease Using TemplateMatching And Support Vector Machine Algorithms”, FirstInternational Symposium On Computing And Networking,2013.

[13] L. Lucchese and S.K. Mitra, „Color Image Segmentation: AState of-the-Art Survey‟, Proceeding of the Indian NationalScience Academy, Vol. 67A, No. 2, 2001, pp. 207-221.

[14] Weizheng S.,Yachun W.,Zhanliang C.& HongdaW.[2008]Grading Method Of Leaf Spot Disease Based OnImage Processing, Proceedings Of 2008 InternationalConference On Computer Science And SoftwareEngineering, Volume 06.

[15] Nithya.A, And Dr.V.Sundaram, Classification RulesFor Indian Rice Diseases, IJCSI International Journal OfComputer Science Issues, Vol. 8, Issue 1, January2011,ISSN (Online): 1694-0814,PN:444-448.

[16]Rajnijain, Sonajharia Minz1 and Ramasubramanian.V, Machine Learning for Forewarning Crop Diseases. J.Ind. Soc. Agril. Statist.63 (1), 2009, PN: 97- 107.

[17] Zhangjian, Zhangwei, Inst. Of Intel. Machines,Chinese Acad. Of Sci., Hefei, China, Support VectorMachine For Recognition Of Cucumber Leaf Diseases,Advanced, and Issue Date: 27-29 March 2010 PN: 264 –266.

[18] J.Savithri, H.Inbarani, “Comparative Analysis Of K-Means, Pso-K-Means, And Hybrid Psogenetic K-Means For Gene Expression Data”, International Journal of Innovations in Scientific and Engineering Research Vol 1 Issue 1 ,PN:43-50,2014.

Rajesh, M., and J. M. Gnanasekar. "Annoyed Realm Outlook Taxonomy Using Twin Transfer Learning." International Journal of Pure and Applied Mathematics 116 (2017): 547-558.

Rajesh, M. & Gnanasekar, J.M. Wireless Pers Commun (2017),https://doi.org/10.1007/s11277-017-4565-9

Rajesh, M., and J. M. Gnanasekar. "GCCover Heterogeneous Wireless Adhoc Networks." Journal of Chemical and Pharmaceutical Sciences (2015): 195-200.

Rajesh, M., and J. M. Gnanasekar. "CONGESTION CONTROL IN HETEROGENEOUS WANET USING FRCC." Journal of Chemical and Pharmaceutical Sciences ISSN 974: 2115.

Rajesh, M., and J. M. Gnanasekar. "GCCover Heterogeneous Wireless Ad hoc Networks." Journal of Chemical and Pharmaceutical Sciences (2015): 195-200.

Rajesh, M., and J. M. Gnanasekar. "CONGESTION CONTROL USING AODV PROTOCOL SCHEME FOR WIRELESS AD-HOC NETWORK." Advances in Computer Science and Engineering 16.1/2 (2016): 19.

RAJESH, M. "TRADITIONAL COURSES INTO ONLINE MOVING STRATEGY." The Online Journal of Distance Education and e-Learning 4.4 (2016).

Rajesh, M. "Object-Oriented Programming and Parallelism."

Rajesh, M., K. Balasubramaniaswamy, and S. Aravindh. "MEBCK from Web using NLP Techniques." Computer Engineering and Intelligent Systems 6.8: 24-26.