Fuzzy-Filtered Neural Network for Rice Disease Diagnosis using Image Analysis

Toran Verma, Sipi Dubey

Abstract: Image mining plays a vital role in the decision-making process in many application areas. Image mining is part of information processing and management. Plant diseases compromise productivity which impacts social life and economy of the nation. The effective use of agriculture image mining can enhance yield production and give economic benefit to the farmer and the country. The research aimed to automate rice diseases identification using image mining for quick diagnosis of the diseases. The digitally captured disease infected and disinfected plant images stored in the database which carries unique feature descriptor in the form of color information, texture appearance, and spatial-frequency information. In this research, the digitally acquired five categories of infected and one category of disinfected images stored in JPEG format in the database. Each category defines unique image features. The acquired images are accessed in RGB color space and cropped and resized in pre-processing steps. All pre-processed images are segmented using Otsu's two-level threshold on a* components of L*a*b* color space image. The segmentation process generates three segments for each image. The 54 hybrid features are extracted using image analysis which includes 6 color entropy, 24 texture, and 24 wavelets F-ratio of spatial-frequency components. The two-way ANOVA analysis is applied in wavelet features to evaluate F-ratio. The extracted features are passed in the CART to select relevant features according to the Gini index split point. The CART created a binary decision tree, reduces 54 attributes to only 13 relevant attributes. The CART selected 13 attributes forwarded in FIS for fuzzy filtering which summarized 13 attributes to 6 attributes. The fuzzy filtered outcomes used to train MLPNN using Scaled Conjugate back-propagation training algorithm to design rice diseases recognition model. The CART feature selection and fuzzy filtering process applied to summarize relevant input features which reduce the complexity of MLPNN. The hybrid CART-FIS-MLPNN model gives 97.1% training and 95.47% testing efficiency.

Keywords: Two-way ANOVA, Classification and Regression Tree (CART), Fuzzy Inference System (FIS), Multilayer Perceptron Neural Network (MLPNN), Image Processing, Pattern Recognition.

I. INTRODUCTION

In 2017, the agriculture sector contributed 6.4% to world GDP growth as per the United Nations report. The agriculture sector employed an approximate 50% workforce and contributed 17-18% of India's GDP in 2017-18. The agriculture sector plays a pivotal role in the Indian economy. Total workforce in the agriculture sector reduced to 25.7 % in 2050 as per estimation. Thus farm mechanization and process automation must be enhanced. The climatic change, conditions of regions and agronomic conditions influence the production of rice in India. The diseases are another factor influencing the production of rice. The diseases may be pathogenic or non-pathogenic causes grain damage and production loss. Plant disease is a dynamic process induced by an incitant (pathogen/abiotic factor) that disturb the energy utilizing the system in plants. The disease may manifest at the micro (biochemical/ physiological/ cytological) and the macro level (symptoms). The diseased plant's performance to produce or survive compromised. Discoloration is one of the most common symptoms observed in a diseased plant, whether due to parasitic or non-parasitic agents. It may reflect a change in the color of the whole plant or one or more of its parts. So Integrated Disease Management (IDM) using information processing and management is monumental to reduce production loss. The timely identification of diseases and correspondingly considered remedial measures can reduce production loss. The identification of plant diseases depends upon color appearance, texture, and shape upon leaves, root, neck, and branches. The different techniques such as immune fluorescence techniques [1], thermography [2][3], fluorescence imaging [4], gas chromatography techniques [5], chain reactions [6] and DNA/RNA based affinity biosensor [7], etc. have been frequently used for quality evaluation of leaves. These methods are inconsistent, inefficient and time taken.

The visual observation based on experience, percept, and intuitive judgment is still the most widely used method for the identification of disease. To identify the plant diseases with the help of a visual process is a hard task as it provides severely low accuracy, takes much time and costly as well. Another disadvantage of diseases identification and classification based visual process is that it can apply in the limited number of places. Automated recognition of plant diseases has got significant research interest in recent years. Sophisticated image processing coupled with advanced computer vision techniques results in accurate and fast identification with less human effort and labour cost. Therefore, automatic, inexpensive, precise technique is required to automate the rice disease recognition and prediction process. The image mining using soft computing techniques, work as a useful tool to automate diseases recognition and prognosis process.

Revised Manuscript Received on May 25, 2019.

Toran Verma, Department of Computer Science & Engineering, Rungta College of Engineering and Technology, Bhilai (C.G.), India.
Sipi Dubey, Department of Computer Science & Engineering Rungta College of Engineering and Technology, Bhilai (C.G.), India.

Generally, computer-based rice diseases recognition composed of three significant steps including pre-processing and segmentation, feature extraction, and recognition. For such systems, there exist many challenges that must be addressed to meet the high detection and classification accuracy. The selection of unique feature descriptor for each disease and selection of efficient models are the major challenge for accurate disease identification. The prominent feature selection and correspondingly, method implementation improve automated diseases identification.

This research aims to automate rice disease recognition process, for accurate and on-the-spot diagnosis. This work proposes an automatic system for the identification of rice plant diseases using image mining of discoloured features of rice plant diseases and soft computing model. The proposed system consists of four significant steps, i.e., image pre-processing and segmentation for the region of interest (ROI) selection, hybrid feature extraction, Feature summarization, and disease identification model implementation. At first, image pre-processing and segmentation performed on the original images captured in the uncontrolled environment to extract diseased part of images. Secondary, color entropy, texture, spatial-frequency wavelet features are extracted to create a unique feature descriptor for each category of rice plant image. In the next step, CART and FIS are used for feature summarization and finally, MLPNN model applied for automated disease identification. The significant contributions of this work are listed as follows:

- The spatial-frequency wavelet F-ratio feature descriptor of the segmented image using two-way ANOVA analysis has done. It reflects within and between covariance features between segments of each image. The combination of wavelet F-ratio with texture and color entropy features makes unique feature descriptor for the six categories of images.
- The CART-FIS-MLPNN hybrid model implemented for rice disease recognition. The CART applied for relevant feature selection, and FIS implemented for filtering of extracted hybrid features. These optimized features used in MLPNN for the disease recognition automation process.

The remaining paper structured in the following manner: The literature survey done in Section 2. The theoretical analysis of the methods explained in Section 3. The proposed methodology discussed in Section 4 and the corresponding experimental results depicted in Section 5. Finally, the research findings concluded in Section 5.

II. LITERATURE REVIEW

Integrated Plant Disease Management (IPDM) is a combination of human, information processing, and management. It mainly deals with decision making for plant diseases identification and management. The image processing based machine learning performs a vital task in the implementation of IPDM. The visible symptoms in leaves and panicle considered in this research. Most diseases breed some demonstration in the visible spectrum. In general cases, the diagnosis or guess about the disease is performed visually by humans by their expertise, but dynamic changes in visible symptoms due to dynamic behaviour of environment caused inaccuracy in judgment, counter-effect is the significant loss in production. Therefore, researchers have exerted to computerize the method of plant disease recognition and categorization using technologies.

The visible patterns of rice plant diseases appear in the various fraction of a plant like a panicle, stems, leaves, and roots [8]. White to gray-green spots with green border appears in leaf blast disease in the initial stage. The spindle-shaped, whitish graycenters appear in the earlier stage. The observed symptom of panicle blast is the greyish brown lesion in the neck. Small, circular, yellow-brown or brown lesions observed on the leaves in brown spot. Greenish-gray oval-shaped lesions appear on the leaf sheath in sheath blight. The observed symptoms of stem borer are dead tillers in vegetative level and whiteheads during the generative stage [9]. Digital still cameras (DSCs) with inbuilt image processing algorithm are frequently used for digital image acquisition [10]. The colors represented in the encoded format [11]. Image processing surrounds analog or digital signal processing to get an improved image or to mine image feature descriptor.

Many investigators worked on the various types of plant images for image mining purpose. The outcomes of the researches reviewed in the following segments. The review work divided into three segments. The first segment highlights the contribution of investigators in plant diseases recognition using image mining. In the second segment, the applications of image mining to recognize rice plant diseases reviewed. The third segment carries a concluding remark and suggestion given by investigators for plant diseases recognition.

Computer vision with artificial intelligence applied in farming for the decision-making process using image mining [12]. The Pearson Correlation Coefficient of texture features was used in the Naive Bayes classifier to automate Yellow Vein Mosaic Virus (YVMV) disease classification [13]. The apple fruit disease classifier was designed using image analysis and multi-class support vector machine (MSVM) [14]. The method was implemented to identify leaf injury of Clover plant using image feature vector and linear discriminate analysis [15]. The system developed to identify aphid in wheat crop using HOG features digital images and SVM [16]. The method implemented to identify soybean disease using the bag-of-visual-words of local descriptors feature of acquired images, and the SVM technique for three types of soya bean diseases [17]. A hybrid image processing method using candidate hot-spot identification and the statistical inference developed for disease identification in uncultivated condition. The research accomplished in three European wheat diseases [18]. The mobile application developed for leaf diseases diagnosis [19]. The system was implemented to automate the disease identification and disease severity evaluation for six soya bean plant foliar diseases [20].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

A three-layer method was designed to detect white-backed plant-hoppers and their growth stages in paddy using image analysis. The Ada-Boost and SVM classifier used for detection purpose with HOG, LBP and Gabor features [21]. The system was developed to identify leaf disease based on pyramid histogram and wavelet feature, and Random Forest classifier [22]. Advanced CNN architectures such as inceptionV3 are used to identify plant diseases for plant diseases recognition with high accuracy [23]. The requirement of large datasets with varying conditions for accurate classification of plant diseases is the severe limitation of CNN classifier [24].

Image processing based soft computing model is also implemented to identify rice disease. A well articulated reviewed work was done on the application of image mining to recognize rice plant diseases. The review work highlighted the importance of feature selection in the designing process, and the need for advanced mining approach to enhance recognition accuracy [25]. The image processing and rule-based fuzzy system were developed to identify the acuteness of rice plant diseases with 86.35% accuracy [26]. The rice plant images were monitored and analyzed using uncrewed aerial vehicle (UAV) for disease deterrence and insect control [27]. The computer model was developed to foresee rice seed germination using image analysis and machine learning. The color, morphological and textural hybrid feature descriptor is used to design the software model [28]. The fuzzy analysis of images with multi-linear regression procedure used for grade and volume assessment of stored paddy by electronic sensors [29]. The SVM model was designed to identify three types of rice diseases with various textures and shape descriptor. The study recommends for the hybrid feature set to design any model [30]. The SVM model implemented on vegetation indices-based texture features of two diseased rice crop images for automatic detection [31]. K-nearest neighbor (K-NN) and Minimum distance classifier (MDC) implemented on color and shape features of four diseased rice image for identification of disease [32]. The hybrid features applied to design random forest classifier to identify one rice disease [33]. An automated system developed to identify rice plant diseases using the deep convolution neural network. The model was implemented to identify 10 different rice plant diseases. The disease recognition accuracy was 95.48% [34]. The classifier for incremental data sets was implemented using Particle Swarm Optimization and Association Rule Mining. The system optimizes the classification rule generated by associative mining during dynamic changes in datasets. The model reported 84.02% rice disease recognition accuracy [35].

Digital image acquisition, pre-processing, feature extraction, and finally, disease recognition model implementation is common in plant image mining for diseases identification. The significant differences in approaches are the ROI selection, feature extraction, and recognition model implementation. The following gaps have found in the literature survey.

- The models implemented according to extracted features of input images without evaluating relevance and redundancy of input attributes. The model’s performance reduced due to redundant information in the training datasets.
- The model performs well for a limited number of diseased class. When the number of class increases, performance degrades dramatically.
- The high complexity observed in neural network implementation due to noisy unfiltered data. It causes an increase in time and space complexity during training.
- Advanced technology like deep learning based CNN cannot be generalized for all type of images. The performance of deep learning methods is excellent for image data sets captured in the controlled environment. It needs specific hardware requirements with a bulk amount of image data sets. The CNN performs well in images where ROI object boundary is distinct.

In this research, the methodology is proposed to overcome the research gap by the creation of unique image feature descriptor, relevant and filtered feature selection and implementing CART-FIS-MLPNN hybrid model for automated disease recognition system.

III. THEORETICAL ANALYSIS

The image provides information about color, spatial and frequency details which forms the unique pattern. The spatial-frequency components of image is extracted using discrete wavelet transform (DWT) [37] [38] [39]. In DWT, the input decomposed into four lower resolution (or lower scale) components. The DWT characterized by transform kernel pair, represented as three separable 2-D wavelets called horizontal, vertical and diagonal wavelets, respectively and one separable 2-D scaling or approximation function defined by $\varphi(x)$ and $\psi(y)$ which are successive orders of double-resolution representations of themselves, defined as Eq. (1) and Eq. (2).

$$\varphi(x) = \sum_{n} h_{\varphi}(n)\sqrt{2}\varphi(2x-n) \tag{1}$$

$$\psi(x) = \sum_{n} h_{\psi}(n)\sqrt{2}\varphi(2x-n) \tag{2}$$

where h_{φ} is called scaling, and h_{ψ} is called wavelet vectors [37]. The two-dimensional wavelet transforms using hierarchical pyramid method with Haar feature is applied to evaluate wavelet descriptor of the image.

The F-ratio using two-way ANOVA represents the image's spatial-frequency descriptor of approximation, horizontal, vertical and diagonal wavelet components. The two-way ANOVA analysis of wavelet components splits variances for analytical purpose. The spatial-frequency wavelet features are summarized using two-way ANOVA [41] [42].

Retrieval Number: H11050688S319/19©BEIESP
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

These variance features are used to evaluate wavelet coefficient F-ratio. The F-ratio of wavelet features of images defined as (3).

$$F\text{-ratio}=\frac{\text{variance between training sample}}{\text{variance within training sample}} \quad (3)$$

The color entropy H(X) of the image gives the average amount of associated color information in that image [43] defined as (4).

$$H(X) = -\sum_{k=0}^{L-1} p_k \, log \, p_k \quad (4)$$

Where k is intensity ranges $0,1,\dots,L-1$ and p_k is probability of pixels with same intensity range.

The texture feature characterizes the images. It refers to a change of pixel's gray level and color [37] [44]. The gray-level co-occurrence matrix (GLCM) examines texture by analyzing the pixel's spatial relationship. The GLCM transformed images of size $L \times L$ and correspondingly joint probability matrix P_{ij}. Four texture descriptor of rice plant images namely energy, correlation, contrast, and homogeneity extracted which defined as (5), (6), (7) and (8).

$$F_{Energy} = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P_{ij}^2 \quad (5)$$

$$F_{Correlation} = \frac{1}{\sigma_x \sigma_y}\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-\mu_x)(j-\mu_y)P_{ij} \quad (6)$$

where $\mu_x = \sum_{i=0}^{L-1} i \sum_{j=0}^{L-1} P_{ij}$, $\mu_y = \sum_{j=0}^{L-1} j \sum_{j=0}^{L-1} P_{ij}$, $\sigma_x^2 = \sum_{i=0}^{L-1}(i-\mu_x^2)\sum_{j=0}^{L-1} P_{ij}$ and $\sigma_y^2 = \sum_{j=0}^{L-1}(j-\mu_y^2)\sum_{i=0}^{L-1} P_{ij}$

$$F_{Contrast} = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}|i-j|^2 P_{ij} \quad (7)$$

$$F_{Homogeneity} = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P_{ij}/[1+(i-j)^2] \quad (8)$$

The CART extensively used in machine learning, expert system, and multivariate analysis. A CART is a non-linear structure consisting of the internal node for decision making and a labeled leaf node for the characterization of the given data. The CART adopts greedy approach with the highest decrease in impurity (s^*) for decision tree construction in a divide-and-conquer strategy which is defined as (9). The splitting criteria for supervised training data D of an interior node of the binary tree chosen according to heuristic attribute selection [40][45].

$$s^* = maxof\left(G_d - G_{sd\,1}, G_d - G_{sd\,2}, \dots, G_d - G_{sd\,(2^n-2)}\right) \quad (9)$$

Where G_d is Gini index and $G_{sd\,i}$, $1 \le i \le 2^n - 2$ are Gini index of two binary split.

The FIS computing structure composed on the notions of fuzzy logic set theory, if-then rules, and reasoning. The surface structure and deep structure description are two stages of fuzzy modeling. It comprises the identification of input variables, output variables, associated linguistic terms, determination of membership functions (MFs) for the linguistic term, inference rule creation, and defuzzification approach. The subtractive clustering is used to create the number of clusters, cluster center c and spread σ [45]. The fuzzification is done according to evaluated cluster center c and spread σ using the Gaussian membership function $f(x:\sigma,c)$ defined as (10).

$$f(x:\sigma,c) = e^{\frac{-(x-c)^2}{2\sigma^2}} \quad (10)$$

where rice plant feature descriptor is x, the spread of data points is σ, and the cluster center is c. The σ and c are used to fuzzify antecedent part of the fuzzy rules. The rule base for fuzzy filtering depends upon the cluster numbers. The consequent is estimated after the premise part determination using the least-squares method. The Sugeno FIS is used for filtering of inputs to remove redundant data and summarization.

A neural network is a collection of a simple parallel shared processor, with learning capability for use. The synaptic links connect processors according to architectural design. The weight associated with each connection. The assigned weight updated during learning. The SCG back-propagation algorithm used for learning with MLPNN architecture in this investigation [43] [46]. SCG is a variant of Conjugate Gradient method shows super-linear convergence. SCG reduce time complexity by avoiding line-search during learning; perform faster learning as a comparison to other learning techniques. In SCG method, the correction in weight Δw is evaluated according to critical point which reduces global error function $E(w)$ using first order $E'(w)$ and second order $E''(w)$ derivatives of Taylor series expansion. The Δw is defined as (11), where ρ: search direction, and α: step size. The previously assigned weight updated with till stopping criteria satisfied

$$\Delta w = \rho\alpha \quad (11)$$

The rice plant disease recognition efficiency (E) of ANOVA-NARX-KNN based model defined as (12), where CPC is the total number of correctly predicted class of test pattern, and N is the total number of the test pattern.

$$E = \frac{CPC}{N} \times 100 \quad (12)$$

IV. METHODOLOGY

The proposed methodology divided into five sub-sections. The components of the proposed system depicted in **Figure 1**. The digitally acquired images cropped and each image segmented into three segments using Otsu's two-level threshold. The color entropy and texture features of the three segments extracted individually. The co-relational spatial-frequency features between segmented images are extracted using wavelet transform and two-way ANOVA analysis. The all extracted features are combined, normalized and aggregated with desired rice diseased and disinfected categories. The feature sets forwarded in the CART for relevant feature selection. The relevant features with the desired category further filtered using FIS. The filtered features applied in MLPNN for automatic rice diseases categorization.

The critical components of the applied methodology manifested in **Figure 1**. The proposed design is an extension of earlier work [36]. The color images of five diseases infected and one disinfected image of rice crop obtained utilizing a digital camera in JPEG compressed format. The images extracted in the RGB model for cropping and resizing. The cropping performed to get the region of interest (ROI) manually. The ROI images were segmented using Otsu's two-level of the threshold assume l_1 and l_2. The three segments ($SegImg_1$, $SegImg_2$ and SegImg_3) of each ROI of all six categories achieved according to the threshold level less than or equal to l_1, between l_1 and l_2 and greater than l_2. The all three segmented RGB images transformed in L*a*b* color model and color entropy H(X), texture, and spatial-frequency wavelet descriptor using two-way ANOVA calculated for a* and b*. These hybrid descriptors define unique patterns for each category. These three segmented images were forwarded to extract spatial-frequency wavelet F-ratio, color entropy, and texture features. The CART used to select relevant features after feature extraction. The extracted 54 features forwarded in the CART for relevant feature selection. The selected relevant features were filtered using fuzzy filtering. The Sugeno FIS used for fuzzy filtering in this research. The 13 selected relevant features forwarded in Sugeno FIS for filtering. Six fuzzy filtered outcomes of six categories with desired response applied in the multilayer perceptron neural network (MLPNN) using SCG back-propagation (BP) algorithm for rice plant disease recognition.

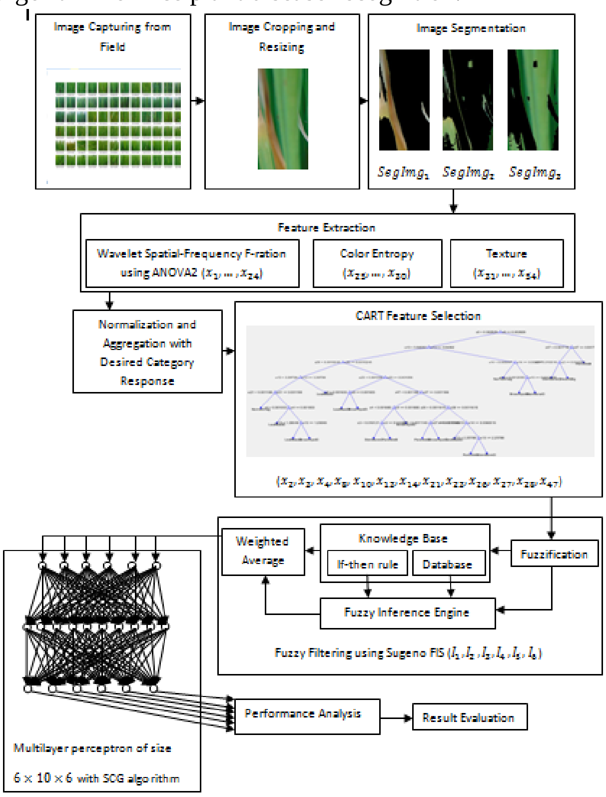


Figure 1: Components of automated rice diseases recognition model

The methodology steps are summarized in following pseudo code.

// Methodology Pseudo Code//

1. *Capture image IM from the rice crop field using digital camera*
2. ***For** each captured image **do***
   begin
   a. *Image pre-processing; cropping and resizing*
   b. *Create three segments using Otsu's multi-level threshold method*
   c. *Extract features for three segments*
      i. *Wavelet spatial-frequency F-ratio using Two-way ANOVA analysis*
      ii. *Color entropy features*
      iii. *Texture features*
   end
3. *Normalize extracted features and aggregate with desired categories*
4. *Apply binary CART for feature selection*
5. *Apply Sugeno FIS for feature Summarization*
6. *Divide datasets as train and test data*
7. *Create MLPNN model with SCG learning law*
8. *Train MLPNN with train data*
9. *Evaluate Performance of MLPNN using test data*

V. RESULT AND DISCUSSION

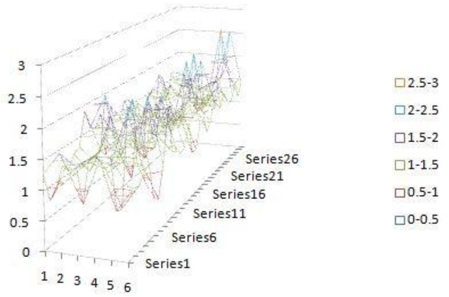
The diseases infected, and disinfected images of rice plant captured in JPEG (JPG) format with the dimension of 5152×3864 from the rice field in day-lighting in the uncontrolled environment by digital camera SONY/DSC-H300, and stored in the database. The entire experiment conducted in MATLAB Version: 8.4.0.150421 with machine configuration; Intel(R) Pentium(R) CPU N3530@ 2.166 GHz processor, 2 GB RAM, 500 GB HDD and Windows 8 Pro 64-bit operating system. The stored images were accessed in RGB color space, cropped, and resized at 205×410. The resized images carrying most diseased parts, manually selected for segmentation. In this research, six categories of rice plant images with 30 images with each category had considered. The five categories belonged to diseased rice plant, and one category was a disinfected image. The considered infected rice plants were the Brown spot, Panicle blast, Leaf blast, Stem Borer and Sheath blight. The snapshot of the pre-processed panicle blast infected 30 images is depicted in **Figure 2**.

Retrieval Number: H11050688S319/19©BEIESP
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

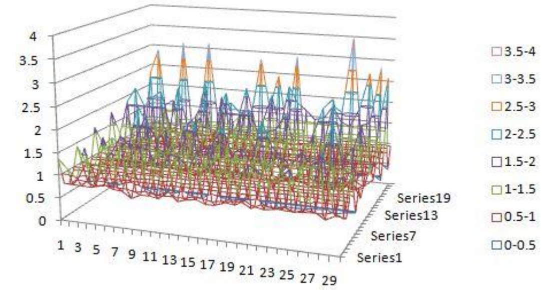
Figure 2: Snapshot of 30 cropped panicle blast infected images

The selected images transformed from RGB to L*a*b* color space. The Otsu's two-level global threshold technique applied in a* components of L*a*b* color images. According to the evaluated threshold value of a* component, the quantization and binarization operation performed. The created three binary matrices were applied in RGB images to create three segments of each image [38].

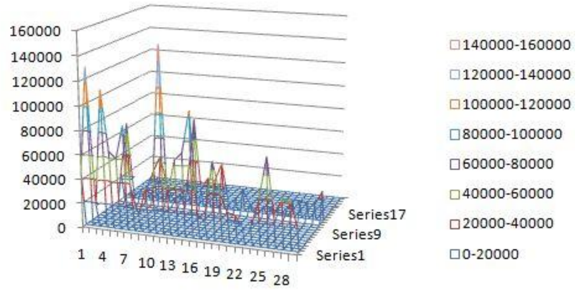
The 24 spatial-frequency F-ratio using two-dimensional ANOVA, 6 color entropy, and 24 texture features of the segmented image of all six categories extracted. **Figure 4** depicts the extracted wavelet F-ratio, color entropy, and texture features of brown spot infected images of the rice crop in the 3D graph. The features of the remaining five categories extracted in the same way. The features were aggregated to create unique 54 hybrid feature descriptor (x_1, $x_2, \ldots, x_{54}$) of each image of six categories. The aggregated features associated with distinct desired class for six categories before processing. Each category name considered as desired classes in the CART. The CART applied on hybrid feature sets to extract relevant features. **Figure 5** depicts the outcome of the CART method for relevant feature selection. The attributes appeared in branches of CART reflect the relevant features as (x_2, x_3, x_4, x_8, x_{10}, x_{13}, x_{14}, x_{21}, x_{23}, x_{26}, x_{27}, x_{28}, x_{47}). The CART model selected 13 features out of 54 features as relevant.



(a)



(b)



(c)

Figure 4: Panicle blast features for L*a*b* color transformed, 30 input images (a) Entropy features (b) Texture feature (d) DWT F-ratio

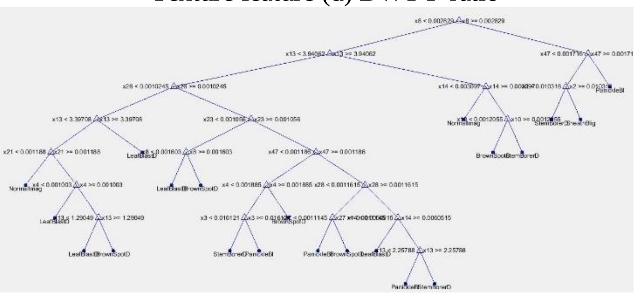


Figure 5: CART outcome with relevant attributes ($x_2, x_3, x_4, x_8, x_{10}, x_{13}, x_{14}, x_{21}, x_{23}, x_{26}, x_{27}, x_{28}, x_{47}$)

The CART selected features (x_2, x_3, x_4, x_8, x_{10}, x_{13}, x_{14}, x_{21}, x_{23}, x_{26}, x_{27}, x_{28}, x_{47}) with desired response were applied in Sugeno FIS model. The assumed desired response for six categories were [1 0 0 0 0 0], [0 1 0 0 0 0], [0 0 1 0 0 0], [0 0 0 1 0 0], [0 0 0 0 1 0] and [0 0 0 0 0 1]. The simulated result of Sugeno FIS model with 13 inputs and 6 outputs ($I_1, I_2, I_3, I_4, I_5, I_6$) is depicts in **Figure 6.**

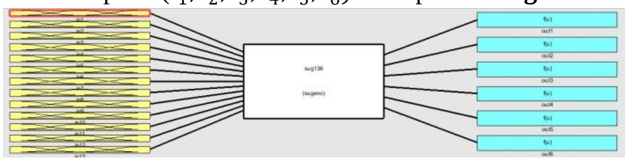


Figure 6: Sugeno FIS with 13 inputs and 6 outputs

Total seven if-then rules formed in Sugeno FIS for feature filtering. The seven Sugeno FIS rules applied on 13 CART selected relevant features. The fuzzy filtering process reduced the 13 input features into 6 features ($I_1, I_2, I_3, I_4, I_5, I_6$) by data reduction.

The filtered feature set with the desired response of six categories used to design MLPNN model using SCG back propagation algorithm. The datasets divided in training and testing sets. The MLPNN trained and tested, ten times with different initialization of weights. The one sample instance using MATLAB is shown in Figure 8,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Figure 9, Figure 10, Figure 11, Figure 12 and Figure 13. **Figure 8** depicts MLPNN architecture with 6 input, 10 hidden and 6 output neurons. The 6 inputs $(I_1, I_2, I_3, I_4, I_5, I_6)$ were the filtered outcome of Sugeno FIS. The 6 output neuron were considered according to six categories used in this research with desired responses for categories were[1 0 0 0 0 0], [0 1 0 0 0 0], [0 0 1 0 0 0], [0 0 0 1 0 0], [0 0 0 0 1 0] and [0 0 0 0 0 1] respectively.

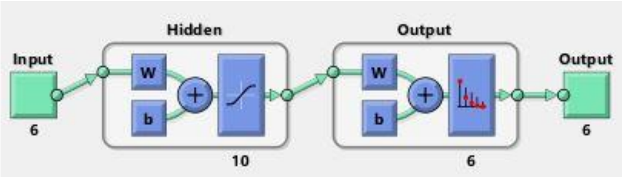


Figure 8: BPNN Architecture with size $6 \times 10 \times 6$

The performance graph of one instance represented in **Figure 9**, and training state for validation parameter shown in **Figure 10**. **Figure 11** depicts error histogram, and **Figure 12** depicts the region of the curve which shows training and validation error. The overall performances are displayed by the confusion matrix in **Figure 13**. The error histogram, region of curve and confusion matrix shows minimal training and validation error, and it can accurately test random pattern.

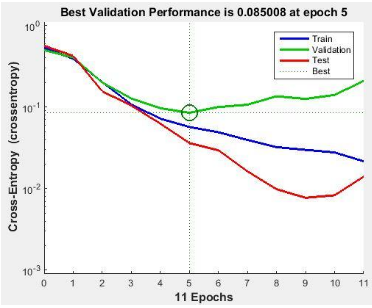


Figure 9: Performance graph

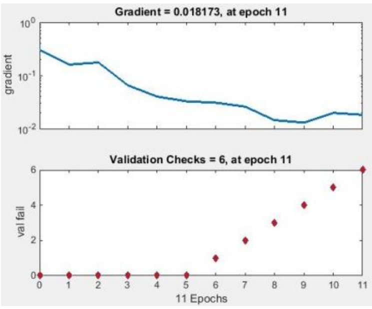


Figure 10: Training state

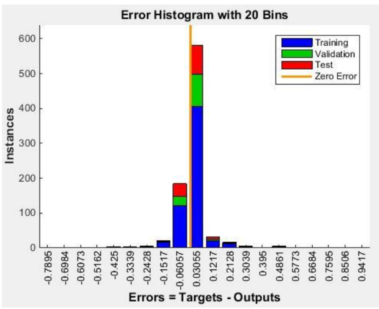


Figure 11: Error histogram

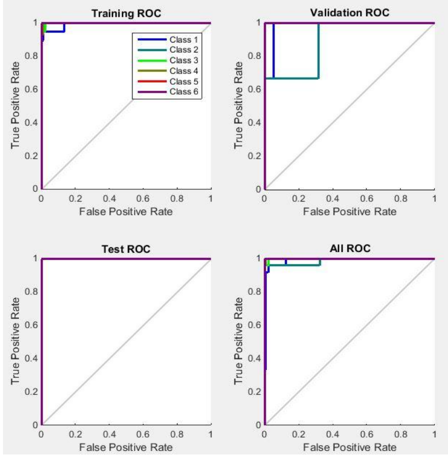


Figure 12: Region of curve

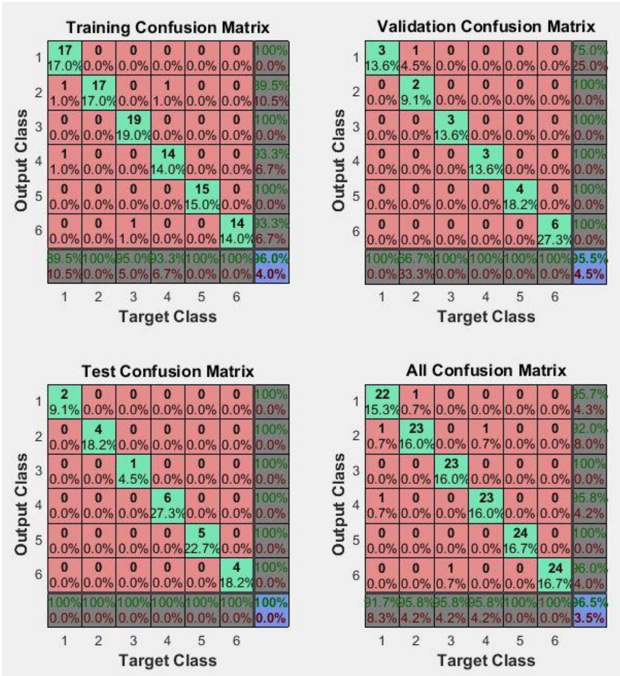


Figure 13: Confusion matrix

The BPNN algorithm executed 10 times to check the robustness of the method. Initially, all 54 extracted features applied in BPNN without features selection and fuzzy filtering. The average performance of ten simulation results with $54 \times 10 \times 6$ MLPNN is 63.8% for training, 56% for validation and 59.23% for testing. The CART selected 13 features without fuzzy filtering applied in $13 \times 10 \times 6$ MLPNN, and the average performance is 61% for training, 59% for validation and 63% for testing. The proposed method; CART features selected and FIS fuzzy filtered output applied in $6 \times 10 \times 6$ MLPNN and the average performance is 97.1% for training, 95% for validation and 95.47% for testing.

The method was evaluated with other different possibilities to perform rice diseases pattern recognition. The summary reports of training and test performance depicted in **Table 5**. Table 5 summarizes the performance of CART, FIS, MLPNN and hybrid structure as rice pattern disease recognizer.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

The CART-FIS-MLPN architecture gives the best performance regarding training and testing performance.

Table 5: Comparative performance of different models and its hybrid structure

Model Number	Model Name	Training Performance (%)	Testing Performance (%)
1	CART	88.89	63.88
2	FIS	100	25
3	MLPNN	63.8	54.63
4	CART	88.89	63.88
5	CART-FIS	85.395	51.8475
6	CART-BPNN	61.1	63.19
7	CART-BPNN-FIS	90.97	66.67
8**	**CART-FIS-BPNN****	**97.1**	**95.47**

The experimental results compared with five other similar work. The images captured with the help of the high-resolution camera. Since the captured images are large, some of the authors [31] [32] [33] reduce the size of the images by cropping. The disease detection divided into two steps: (i) First step consists of image collection, preprocessing, segmentation, feature extraction, and (ii) the second step consists of detection of diseases. The result of the proposed method also compared with CNN based model, where images compressed instead of cropping and low-level variance features extracted for CNN model implementation [36]. The comparative results depicted in **Table 6** according to rice diseases detection accuracy.

Table 6: Performance Comparison

Author name and Year	Method	Number of diseased class and disinfected class	Features	Average Accuracy
SantanuPhadikar et al. [31]	SVM	2 (diseased)	Vegetation indices-based texture features	84%
Amrita A. Joshi et al. [32]	K-nearest neighbor (K-NN) and Minimum distance classifier (MDC)	4 (diseased)	Color and shape features	K-NN: 87.02 % MDC: 89.23 %
Xiaochun Mai et al. [33]	Random forest Classifier	1(diseased)	Color, shape and texture features	88.5%
Yang Lu et al. [34]	Deep convolution neural networks	10 (diseased)	low-level variance features	95.48 %
Proposed Methodology*	**CART-FIS-BPNN**	**5 (diseased) 1 (disinfected)**	**Wavelet F-ration, Color Entropy, Texture**	**95.47 %**

The comparative results depict that proposed method give better performance to recognize rice plant diseases comparison to [31] [32] [33], and it can also consider an alternative method for [34] with reduced feature sets. Hence, the proposed method and hybrid feature descriptor legitimate the desired objective.

Rice disease diagnosis using image mining can be used by farmers and government agencies to identify and quantify the presence of pathogens, assess the effectiveness of protection technologies, and estimate the yield loss. The pesticide industry can use this approach to certify planting materials and to detect new pathogens.

VI. CONCLUSION

The timely diagnosis of crop diseases is essential to reduce production loss and pest management. Therefore the image processing based intelligent system is designed to perform rice crop disease recognition. The new hybrid feature descriptor has created using color entropy, spatial-frequency wavelet F-ratio using two-way ANOVA and texture components of three segmented images of each input images. The CART has used to select relevant features from extracted hybrid 54 features. CART generates 13 relevant features for 54 input feature sets. Fuzzy filtering process removes redundant data from the dataset and performs data summarization. Sugeno FIS summarizes CART generated 13 features in 6 features. The feature summarization reduces the design complexity of MLPNN architecture from 54×10×6 architecture to 6×10×6. It also enhances the disease pattern recognition efficiency of MLPNN. The MLPNN is converged quickly and generalized in less number of training feature sets.

The initially extracted 54 feature set of six categories passed in MLPNN, and the average performance of training and testing is 63.8% and 54.63%. The next level of MLPNN performance evaluated with the CART selected 13 relevant feature sets, and training and testing performance is 61.1% and 63.19%. The result reflects a small improvement in test performance. However still due to data redundancy, the MLPNN model not able to generalize and converge.

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication

Therefore CART selected feature sets are forwarded in Sugeno FIS to filter and summarization. FIS generates 6 fuzzy filtered outcomes for each CART selected 13 features using seven if-then rules. These outcomes forwarded in the MLPNN with desired class, and the training and testing performance of the MLPNN is 97.1% and 95.47%. The result reflects radical improvement in rice diseases pattern recognition using CART-FIS-BPNN hybrid method.

The relevant feature selection and summarization perform an essential role in convergence and generalization of any model for pattern recognition. The proposed process can further be improved by enhancing optimal feature selection using the hybrid model of the CART, and FIS with the nature-inspired and genetic algorithm. The improvement in input data sets of the model will automatically enhance the performance of the MLPNN model. It will decrease the complexity of the MLPNN design regarding learning and generalization.

ACKNOWLEDGMENT

We are grateful to Mr. Nalin Lunia (Secretary) and Dr.K.S.Pandya (Principal), Chhattisgarh Agriculture college, Durg, CG (India) for their kind support in this research to collect digital image data sets of rice plant under their esteemed guidance. We are also grateful to Dr.SourbhRungta (Vice Chairman), Mr. SonalRungta (Secretary) and Dr.S.P.Dubey (Dean) for their kind support in this research.

REFERENCES

1. Abouhussein, D.M.N., Sabry, P., Moghawry, H.M.Recent (2019) “insight into critical concerns in international quality regulations governing the pharmaceutical industry: A review” International Journal of Pharmaceutical Research, 11 (1), pp. 708-718.
2. Chaerle, L., Leinonen, I., Jones, H.G. and Van Der Straeten, D., 2006. Monitoring and screening plant populations with combined thermal and chlorophyll fluorescence imaging. Journal of experimental botany, 58(4), pp.773-784.
3. Mahlein, A.K., 2016. Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. Plant Disease, 100(2), pp.241-251.
4. Kuckenberg, J., Tartachnyk, I. and Noga, G., 2009. Temporal and spatial changes of chlorophyll fluorescence as a basis for early and precise detection of leaf rust and powdery mildew infections in wheat leaves. Precision Agriculture, 10(1), pp.34-44.
5. Fang, Y., Umasankar, Y. and Ramasamy, R.P., 2014. Electrochemical detection of p-ethylguaiacol, a fungi infected fruit volatile using metal oxide nanoparticles. Analyst, 139(15), pp.3804-3810.
6. Cai, H.Y., Caswell, J.L. and Prescott, J.F., 2014. Nonculture molecular techniques for diagnosis of bacterial disease in animals: a diagnostic laboratory perspective. Veterinary pathology, 51(2), pp.341-350.
7. Eun, A.J.C., Huang, L., Chew, F.T., Li, S.F.Y. and Wong, S.M., 2002. Detection of two orchid viruses using quartz crystal microbalance (QCM) immunosensors. Journal of Virological Methods, 99(1-2), pp.71-79.
8. Chaube, H.S. and Pundhir, V.S., 2005. Crop diseases and their management. PHI Learning Pvt.Ltd..
9. Rangaswami, G. and Mahadevan, A., 1998. Diseases of crop plants in India. PHI Learning Pvt.Ltd..
10. Balasubramanian, R., 2003. Device characterization. Digital color imaging handbook, pp.269-379, Chapter 5.
11. Giorgianni, E.J., Madden, T.E. and Splauding, K.E., 2003. Color Management for digital imaging systems-Digital Color Imaging Handbook, Chapter 4.
12. Patrício, D.I. and Rieder, R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. Computers and Electronics in Agriculture, 153, pp.69-81.
13. Mondal, D., Kole, D.K. and Roy, K., 2017. Gradation of yellow mosaic virus disease of okra and bitter gourd based on entropy based binning and Naive Bayes classifier after identification of leaves. Computers and Electronics in Agriculture, 142, pp.485-493.
14. Dubey, S.R. and Jalal, A.S., 2016. Apple disease classification using color, texture and shape features from images. Signal, Image and Video Processing, 10(5), pp.819-826.
15. Kruse, O.M.O., Prats-Montalbán, J.M., Indahl, U.G., Kvaal, K., Ferrer, A. and Futsaether, C.M., 2014. Pixel classification methods for identifying and quantifying leaf surface injury from digital images. Computers and electronics in Agriculture, 108, pp.155-165.
16. Liu, T., Chen, W., Wu, W., Sun, C., Guo, W. and Zhu, X., 2016. Detection of aphids in wheat fields using a computer vision technique. Biosystems Engineering, 141, pp.82-93.
17. Pires, R.D.L., Gonçalves, D.N., Oruê, J.P.M., Kanashiro, W.E.S., Rodrigues Jr, J.F., Machado, B.B. and Gonçalves, W.N., 2016. Local descriptors for soybean disease recognition. Computers and Electronics in Agriculture, 125, pp.48-55.
18. Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A.D. and Ortiz-Barredo, A., 2017. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. Computers and Electronics in Agriculture, 138, pp.200-209.
19. Prasad, S., Peddoju, S.K. and Ghosh, D., 2016. Multi-resolution mobile vision system for plant leaf disease diagnosis. Signal, Image and Video Processing, 10(2), pp.379-388.
20. Shrivastava, S., Singh, S.K. and Hooda, D.S., 2015. Color sensing and image processing-based automatic soybean plant foliar disease severity detection and estimation. Multimedia Tools and Applications, 74(24), pp.11467-11484.
21. Qing, Y.A.O., Chen, G.T., Zheng, W.A.N.G., Zhang, C., Yang, B.J. and Jian, T.A.N.G., 2017. Automated detection and identification of white-backed planthoppers in paddy fields using image processing. Journal of Integrative Agriculture, 16(7), pp.1547-1557.
22. Ahmed,A.A.N., Haque, H.F., Rahman, A., Ashraf, M.S. and Shatabda, S., 2019. Wavelet and Pyramid Histogram Features for Image-Based Leaf Detection. In Emerging Technologies in Data Mining and Information Security (pp. 269-278). Springer, Singapore.
23. Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S., Boukhalfa, K. and Moussaoui, A., 2018. Deep Learning for Plant Diseases: Detection and Saliency Map Visualisation. In Human and Machine Learning (pp. 93-117). Springer, Cham.
24. Barbedo, J.G.A., 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and Electronics in Agriculture, 153, pp.46-53.
25. Bera, T., Das, A., Sil, J. and Das, A.K., 2019. A Survey on Rice Plant Disease Identification Using Image Processing and Data Mining Techniques. In Emerging Technologies in Data Mining and Information Security (pp. 365-376). Springer, Singapore.
26. Sethy, P.K., Negi, B., Barpanda, N.K., Behera, S.K. and Rath, A.K., 2018. Measurement of Disease Severity of Rice Crop Using Machine Learning and Computational Intelligence. In Cognitive Science and Artificial Intelligence (pp. 1-11). Springer, Singapore.
27. Li, Y., Qian, M., Liu, P., Cai, Q., Li, X., Guo, J., Yan, H., Yu, F., Yuan, K., Yu, J. and Qin, L., 2018. The recognition of rice images by UAV based on capsule network. Cluster Computing, pp.1-10.
28. Lurstwut, B. and Pornpanomchai, C., 2017. Image analysis based on color, shape and texture for rice seed (Oryza sativa L.) germination evaluation. Agriculture and Natural Resources, 51(5), pp.383-389.
29. Baskar, C., Nesakumar, N., Rayappan, J.B.B. and Doraipandian, M., 2017. A framework for analysing E-Nose data based on fuzzy set multiple linear regression: Paddy quality assessment. Sensors and Actuators A: Physical, 267, pp.200-209.
30. Yao, Q., Guan, Z., Zhou, Y., Tang, J., Hu, Y. and Yang, B., 2009, May. Application of support vector machine for detecting rice diseases using shape and color texture features. In Engineering Computation, 2009. ICEC'09. International Conference on (pp. 79-83). IEEE.

31. Phadikar, S. and Goswami, J., 2016, March. Vegetation indices based segmentation for automatic classification of brown spot and blast diseases of rice. In Recent Advances in Information Technology (RAIT), 2016 3rd International Conference on (pp. 284-289). IEEE.
32. Joshi, A.A. and Jadhav, B.D., 2016, December. Monitoring and controlling rice diseases using image processing techniques. In Computing, Analytics and Security Trends (CAST), International Conference on (pp. 471-476). IEEE.
33. Mai, X. and Meng, M.Q.H., 2016, June. Automatic lesion segmentation from rice leaf blast field images based on random forest. In Real-time Computing and Robotics (RCAR), IEEE International Conference on (pp. 255-259). IEEE.
34. Lu, Y., Yi, S., Zeng, N., Liu, Y. and Zhang, Y., 2017. Identification of rice diseases using deep convolutional neural networks. Neurocomputing, 267, pp.378-384.
35. Sengupta, S. and Das, A.K., 2017. Particle Swarm Optimization based incremental classifier design for rice disease prediction. Computers and Electronics in Agriculture, 140, pp.443-451.
36. Verma, T., Dubey, S. and Sabrol, H., 2017, October. Color image segmentation of disease infected plant images captured in an uncontrolled environment. In International Conference on Next Generation Computing Technologies (pp. 790-804). Springer, Singapore.
37. Gonzalez, R.C., Eddins, S.L. and Woods, R.E., 2004. Digital Image Processing Using MATLAB. Prentice Hall.
38. [38] Balamurugan, V. and Anandhakumar, P., 2009. Neuro-fuzzy based clustering approach for content based image retrieval using 2D-wavelet transform. International Journal of Recent Trends in Engineering, 1(1), pp.418-424.
39. Gu, X., Du, J.X. and Wang, X.F., 2005, August. Leaf recognition based on the combination of wavelet transform and gaussian interpolation. In International Conference on Intelligent Computing (pp. 253-262). Springer, Berlin, Heidelberg.
40. Han, J., Pei, J. and Kamber, M., 2011. Data mining: concepts and techniques. Elsevier.
41. Spiegel, M.R., Schiller, J.J., Srinivasan, R.A. and LeVan, M., 2009. Probability and statistics (Vol. 2). New York: Mcgraw-hill.
42. [42] Kothari, C.R., 2004. Research methodology: Methods and techniques. New Age International.
43. Haykin, S., 1994. Neural networks: a comprehensive foundation. Prentice Hall PTR.
44. Gao, R., Liu, G. and Si, Y., 2010, September. A recognition method of apples based on texture features and em algorithm. In World Automation Congress (WAC), 2010 (pp. 225-229). IEEE.
45. Jang, J.S.R., Sun, C.T., Mizutani, E., 1997. A Computational Approach to Learning and Machine Intelligence.–Upper Saddle River, NJ: Prentice Hall, Inc.
46. Moller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural networks, 6(4), pp.525-533.