

Detection And Classification of the Medicinal Plants for Various Life Threatening Disease Using Recurrent Neural Network

Dr.A.Nithya Rani¹, R.Mohamed Muneerul Subiyan², Mohammed Salih Hemd Mohammed³, S. Prashandhini⁴, R. Ragavan⁵

Associate Professor¹, PG student^{2,3, 4,5}

Department of Information Technology^{1,2,3,4,5}

CMS College of Science and Commerce (Autonomous), Coimbatore, India^{1,2,3,4,5}

ARTICLE INFO

Received: 28 Oct 2024

Revised: 18 Nov 2024

Accepted: 26 Dec 2024

ABSTRACT

Many plant species across the world contains bioactive chemicals which can be used for many life threatening diseases. Many scientists usually conduct many studies and experiment in their paramedical laboratories using plant species. Especially many plant species which is utilized for life threatening disease like cancer has been found across the dense forest and oceans. However on basis of those studies, acquiring of those plant species becomes complex and leads to extraction of incorrect plant species. On advancement of the Internet of Things and artificial intelligence, it becomes feasible to acquire those plant species accurately on employing those techniques for detection and classification task in dense forest and oceans. Traditionally many machine learning and deep learning architecture has been employed to achieve above specified function. Despite of several advantages, those approaches still lags in performance aspects. In order to enhance performance of detection and classification, a deep learning architecture is proposed in this article. Recurrent Neural Network is designed and implemented to detect and classify the medical plants acquired in form of images. Recurrent Neural Network is highly capable of the processing images containing complex structures as it can produce good recognition accuracy. Initially image preprocessing step composed of contrast enhancement is performed to enhance image pixel quality and normalize the complex structure in the image. Preprocessed image is applied to recurrent neural network composed of recurrent layer to identify complex relationship of the image pixel as it is represented in form feature map. Further dense layer in introduced with softmax function to process the feature map to classify and predict the medical plant species. Experimental results of the model are obtained using benchmark Mendeley dataset. On experimental and performance analysis of the model, it is proved that proposed model produces 98.5% recognition accuracy compared to conventional approaches.

Keywords: Medicinal Plants, Recurrent Neural Network, Mendeley Dataset, Deep Learning

1. INTRODUCTION

In ancient times, Medicinal Plants were used for treatment of the many life threatening diseases. Although usage of the medicinal plant species has been decreased as modern medicine utilized but its significance is not reduced. Thus identification of the medicinal plant species to treat life threatening diseases is gaining more significance across the world. However manual observation of the medicinal plant is prone to inaccurate selection and utilization. Thus, it becomes mandatory to employ an automated method using computer vision and artificial intelligence approaches which assist botanist in identification of the plant accurately in dense forest and oceans. Traditionally many machine learning and deep learning architecture has been employed to achieve it. Despite of several advantages, those approaches still lags in performance aspects as it found to be high complex in identifying plant species effected with virus and bacterial diseases.

In order to enhance performance of detection and classification, a deep learning architecture is proposed in this article. Recurrent Neural Network is designed and implemented to detect and classify the medical plants acquired

in form of images. Recurrent Neural Network is highly capable of the processing images containing complex structures as it can produce good recognition accuracy. Initially image preprocessing step composed of contrast enhancement is performed to enhance image pixel quality and normalize the complex structure in the image. Preprocessed image is applied to recurrent neural network composed of recurrent layer to identify complex relationship of the image pixel as it is represented in form feature map. Further dense layer is introduced with softmax function to process the feature map to classify and predict the medical plant species

The remaining part of the article is organized as follows; section 2 provides the review of related works in predicting and classifying plants using machine learning and deep learning architectures. Section 3 defines a new design of the recurrent neural network to perform prediction and classification of the medicinal plants. Section 4 provides experimental and performance results of the proposed model against conventional approaches on processing benchmark dataset. Finally section 5 concludes the article with major findings.

2. RELATED WORK

In this section, review of the related work towards prediction and classification of medical plant using machine learning and deep learning is performed on its design and performance aspects is as follows

2.1. Medicinal plant classification using Random Forest Classifier

In this architecture, machine learning classifier named as Random Forest classifier is employed for medical plant classification. Initially image preprocessing is performed to obtain the high resolution image. Those preprocessed image is employed to feature extraction technique to extract the morphological features such as area, perimeter, length and width. Further those features were employed to the classifier function. Classifier function build decision tree using subset of features. Decision tree represents the classes of medicinal plant.

2.2. Medicinal plant classification using Convolution Neural Network

In this architecture, deep learning model named as Convolution Neural Network is employed for medical plant classification. Initially image preprocessing is performed to obtain the high resolution image. Those preprocessed image is employed to convolution layer of the model to extract the morphological features such as area, perimeter, length and width of the plant image using kernel function. Further those features were employed to the classifier function in fully connected layer. Classifier function uses a decision tree process on generating the subset of features. Decision tree generate the medicinal plant class.

3. Proposed Model

In this section, design of deep learning model using recurrent neural network with long short term memory model along preprocessing process such as contrast enhancement is performed for medicinal plant prediction and classification. Component of the model is as follows

3.1. Image Pre-processing- Contrast Limited Adaptive Histogram Equalization

Contrast limited adaptive histogram equalization is applied as image preprocessing to enhance the resolution of the image pixel properties on contrast enhancement and image normalization. Particular process transforms the image into histogram for each tile. Next Clip limit is used to define maximum allowed height of the histogram for single histogram bin. It is calculated through tile size parameter. Cumulative distribution function and bilinear interpolation is performed on basis of the clipped histogram. Cumulative distribution function provides new pixel intensity mapping to specified tile.

3.2. Recurrent Neural Network

Recurrent Neural Network which uses the multiple layer such as Input layer, hidden layer using long short term memory mechanism and dense layer. Working of each layer is as follows

3.2.1. Input layer

Input layer of the Recurrent Neural Network obtains the pixel information of the image and transforms it into sequential data. Transformed sequential data vector composed of feature representation is projected to hidden

layer of the model. Figure 1 represents the proposed architecture of the medicinal plant classification deep learning model.

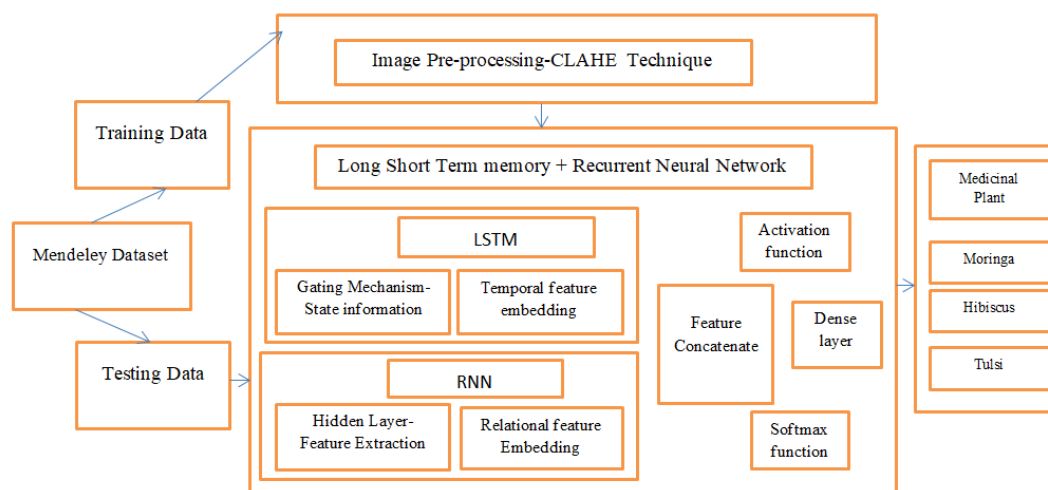


Figure 1: Medicinal Plant Classification Deep learning model

3.2.2. Hidden Layer

Hidden layer of the recurrent neural network processes the image through activation function which uses Pearson correlation coefficients to generate correlated pixel features. Correlated feature in form of state information is projected to LSTM model. Table 1 represents the hyperparameter setting of the Recurrent Neural Network for optimal sequence length. Learning rate, batch size and epoch were employed for early stopping to prevent overfitting. Optimizer was used for efficient training and mean square error as loss function

Table 1: Hyperparameter setting of Recurrent Neural Network

Hyperparameter	Value
Learning rate	10^{-6}
Batch Size	25
Epoch	50
Loss function	Cross entropy
Activation function	ReLu
Dropout rate	L2 regularization

3.2.3. Long short term memory mechanism

Long Short Term Model Network is applied to capture long term dependency of the temporal features of the image using gating mechanism and represents in form of cell states. Initially it processes the input sequence over optimized time windows. Further multiple LSTM layers were stacked to learn temporal patterns utilizing forget gate. Especially dense layer is interfaced to transform the LSTM output into feature vector. Table 2 mentions the hyperparameter setting of the Long Short Term Model Network for optimal image features.

Table 2: Hyperparameter setting of LSTM Network

Hyperparameter	Value
Learning rate	10^{-5}
Batch Size	11
Epoch	10
Loss function	Mean Square Error
Optimizer	Adam

Learning rate, batch size and epoch were employed for early stopping to prevent overfitting. Optimizer was used for efficient training and mean square error as loss function

3.2.4. Dense layer

The temporal feature of the image from LSTM model in form of patterns and relational features embedding from RNN in form of relation were concentrated to form a unified feature vector. Further complex relations among the features is learned [14]. Finally dense layer with activation function linearizes the complex relations and softmax function with classifier function classifies the user feedback with class labels of medicinal plant effectively. Algorithm steps of the recurrent neural network with long short term memory is as follows

Algorithm: RNN+LSTM

Input: Mendeley Dataset – Medicinal Plant

Output: Medicinal Plant Classes – {Hibiscus, Tulsi, Moringa, Indica, Chebula}

Process()

Image Preprocessing()

 CLAHE(Image)

 Histogram ()

 Clip Limit (Histogram tile)

 High Contrast _ Preprocessed Image Cumulative distribution function (Clip histogram)

RNN_LSTM ()

 Input layer

 Transform Image into Sequential pixel vector

Hidden Layer ()

 Compute Similarities between features of the Sequential pixel

Relational Embedding

 LSTM ()

 Gating (Association Pattern of Sequence Pixel)

 Cell state = patterns of the sequence

 Feature Vector of LSTM= Long Term Dependency

 Temporal Embedding

Dense layer ()

 Activation function _ReLU (feature map)

 Linear feature map

 Dropout layer_L2 Regularization (linear feature)

 Feature Concatenation (Temporal Feature Vector + Relational Feature Vector)

 Unified Aggregated feature vector

 Softmax function _Classifier (feature Vector)

Class= { Hibiscus, Tulsi, Moringa}

4. Experimental Results

Experimental analysis of the model has been carried out using Mendeley plant dataset (Source: <https://data.mendeley.com/datasets/nnytj2v3n5/1>). Analysis is performed in the python environment with tensor flow functionalities and multiple libraries such as NumPy and Pandas for image manipulation, scikit learn for baseline models and Matplotlib for visualization. Panda's libraries for data processing [15]. Dataset is portioned as 60 percent employed for model training and 40 percent for model testing. Grid search was conducted to identify the optimal hyperparameter for model training to LSTM and RNN.

4.1. Performance analysis

Performance analysis of the model is performed using test data through confusion matrix. Confusion matrix generates the values to the true positive, false negative, true negative and false positive parameters. Those parameter values of the matrix demonstrates strong performance of the LSTM +RNN model in classifying the user

feedback of e learning system on integrating the time series data(processed by LSTM network) and relational data(processed RNN network). Model achieved enhanced prediction accuracy of 9.4 percent. Further precision analysis, recall analysis and accuracy analysis were performed as follows

4.1.1. Precision Analysis

Precision analysis is performed to determine no of the feature correctly classified as specified medicinal plant among the total extracted features. Precision analysis represented on parameter of confusion matrix is as follows

$$\text{Precision} = \frac{TP}{TP+FP}$$

Figure 2 represents precision analysis of the model against conventional approaches. It represents model ability towards detecting and classifying different medicinal plant images.

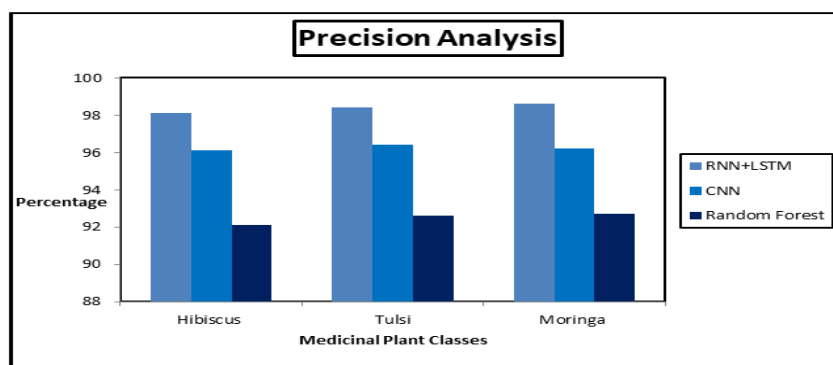


Figure 2: Precision Analysis

4.1.2. Recall Analysis

Recall analysis is performed to determine no of the feature incorrectly classified into specified medicinal plant classes among the total extracted features. Recall analysis represented on parameter of confusion matrix is as follows

$$\text{Recall} = \frac{TN}{TP+FP}$$

Figure 3 represents recall analysis of the model against conventional approaches. It represents model ability towards predicting and classifying different medicinal plants.

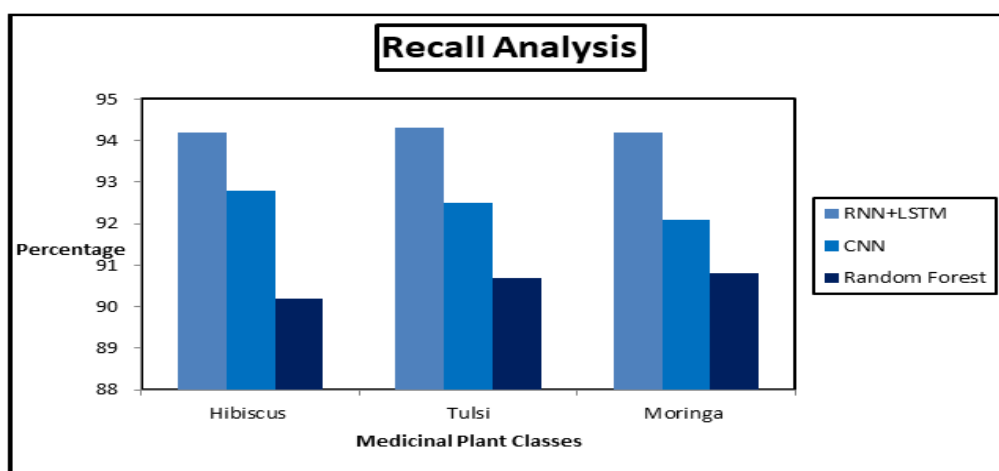


Figure 3: Recall Analysis

4.1.3. Accuracy Analysis

Accuracy analysis is performed as aggregation of recall and precision towards predicting and classifying features into specified medicinal plant classes among total extracted features. F-Measure analysis represented on parameter of confusion matrix is as follows

$$\text{Accuracy} = \frac{TP+TN}{TN+FN+TP+FP}$$

Figure 4 represents Accuracy analysis of the model against conventional approaches. It represents model ability towards predicting and classifying different medicinal plants.

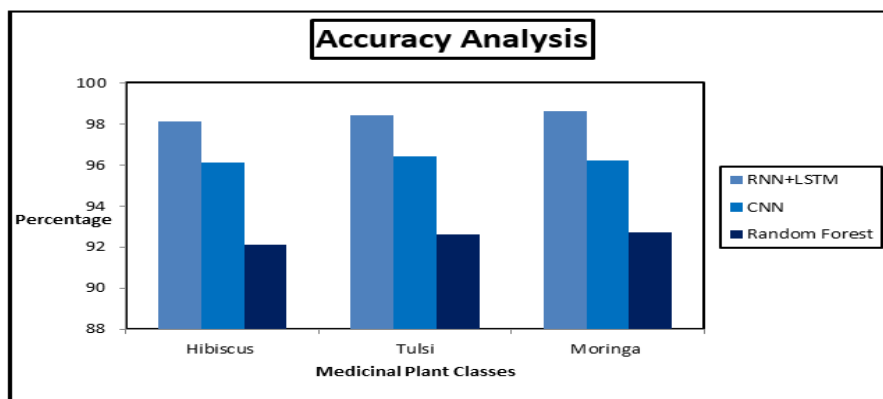


Figure 4: F-Measure analysis

Finally performance of RNN+LSTM architecture performs better with recognition accuracy 98.4% on compared to conventional approaches. Table 2 mentions the performance evaluation of the deep learning architectures in medicinal plant classification.

Hibiscus, Tulsi, Moringa, Indica, Chebula

Table 2: Performance Evaluation of Deep learning architecture in forecasting adoption rate

Technique	Classes	Precision	Recall	Accuracy
RNN+LSTM – Deep learning	Hibiscus	98.1	94.2	98.4
	Tulsi	98.4	94.5	98.2
	Moringa	98.6	94.6	98.1
CNN- Deep learning	Hibiscus	96.4	91.7	96.7
	Tulsi	96.6	91.8	96.3
	Moringa	96.7	92.8	96.7
Random Forest- machine learning	Hibiscus	92.2	90.3	92.8
	Tulsi	92.6	90.4	92.4
	Moringa	92.8	90.3	92.6

CONCLUSION

In this work, Recurrent Neural Network is designed and implemented to detect and classify the medical plants acquired in form of images. Initially image preprocessing is performed using CLAHE technique. Preprocessed image is applied to recurrent neural network to identify complex relationship of the image pixel and those associations are represented in form feature map. Further dense layer is incorporated in the model through softmax function to process the feature map. Decision tree classifier of the softmax function classify feature map into medical plant species classes. Experimental results and performance outcomes of the model proved that proposed model produces 98.5% recognition accuracy compared to conventional approaches

REFERENCES

- [1] Gao, W. and Lin, W., 2012. Frontal Parietal Control Network Regulates the Anti-Correlated Default and Dorsal Attention Networks. *Human Brain Mapping*, 33(1), 192–202.
- [2] Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.X., Chang, Y.F. and Xiang, Q.L., 2007. A Leaf Recognition Algorithm for Plant Classification using Probabilistic Neural Network. 7th IEEE International Symposium on Signal Processing and Information Technology, Giza, Egypt, 11-16.
- [3] Zhang X., Liu Y., Lin H., Liu Y. (2016) Research on SVM Plant Leaf Identification Method Based on CSA. In: Che W. et al. (eds) *Social Computing. ICYCSEE 2016. Communications in Computer and Information Science*, Vol 624, Springer, Singapore.
- [4] Hossain, J. and Amin, M.A., 2010. Leaf Shape Identification Based Plant Biometrics. 13th International Conference on Computer and Information Technology, Dhaka, Bangladesh, 458-463.
- [5] Du, J.X., Wang, X.F. and Zhang, G.J., 2007. Leaf shape based plant species recognition. *Applied Mathematics and Computation*, 185, 883- 893.
- [6] Du, M., Zhang, S. and Wang, H., 2009. Supervised Isomap for Plant Leaf Image Classification. 5th International Conference on Emerging Intelligent Computing Technology and Applications, Ulsan, South Korea, 627-634.
- [7] Herdiyeni, Y. and Wahyuni, N.K.S., 2012. Mobile Application for Indonesian Medicinal Plants Identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram. *International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, West Java, Indonesia, 301-306.
- [8] Prasvita, D.S. and Herdiyeni, Y., 2013. MedLeaf: Mobile Application for Medicinal Plant Identification Based on Leaf Image. *International Journal of Advanced Science, Engineering and Information Technology*, 3, 5–8.
- [9] Le, T.L., Tran, D.T. and Hoang, V.N., 2014. Fully Automatic leaf-based plant identification, application for Vietnamese medicinal plant search. *Fifth Symposium on Information and Communication Technology*, Hanoi, Vietnam, 146-154.
- [10] Arai, K., Abdullah, I.N. and Okumura, H., 2013. Identification of Ornamental Plant Functioned as Medicinal Plant Based on Redundant Discrete Wavelet Transformation. *International Journal of Advanced Research in Artificial Intelligence*, 2(3), 60-64.
- [11] D. K. Berger, "Grey leaf spot disease of maize and food security research," *South Afr. J. Botany*, vol. 100, no. 109, p. 327, Jan. 2017
- [11] C. DeChant, T. Wiesner-Hanks, S. Chen, E. L. Stewart, J. Yosinski, M. A. Gore, R. J. Nelson, and H. Lipson, "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," *Phytopathology*, vol. 107, no. 11, pp. 1426-1432, Nov. 2017.
- [12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards realtime object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137-1149, Jun. 2017.
- [13] P. Xiao, G. K. Venayagamoorthy, and K. A. Corzine, "Combined training of recurrent neural networks with particle swarm optimization and backpropagation algorithms for impedance identification," in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS '07)*, pp. 9–15, April 2007
- [14] K. R. Aravind, P. Raja, K. V. Mukesh, R. Anirudh, R. Ashwin, and C. Szczepanski, "Disease classification in maize crop using bag of features and multiclass support vector machine," in *Proc. 2nd Int. Conf. Inventive Syst. Control (ICISC)*, Jan. 2018, pp. 1191-1196
- [15] Mingjie Lv, Guoxiong Zhou, Mingfang He, Aibin Chen, Wenzhuo Zhang, Yahui Hu "Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet " *IEEE Access*, Vol.8, pp:57952-57966, 2020