



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 8, August 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Medicinal Plant Species Detection using Deep Learning

Mahesh Kumar S M, Mujammil, Naykodi, Pankaj Sharma

Assistant Professor, Department of Computer Science & Engineering, Rajarajeswari College of Engineering, Bangalore,
Karnataka, India

U.G. Student, Department of Computer Science & Engineering, Rajarajeswari College of Engineering, Bangalore,
Karnataka, India

U.G. Student, Department of Computer Science Engineering, Rajarajeswari College of Engineering, Bangalore,
Karnataka, India

U.G. Student, Department of Computer Science & Engineering, Rajarajeswari College of Engineering, Bangalore,
Karnataka, India

ABSTRACT: In the contemporary era, there is a noticeable trend towards automated systems that operate without direct human intervention, owing to rapid advancements in technology. A prime example of this trend is the detection of medicinal plant species, traditionally reliant on human experts with botanical knowledge. In this study, we propose the utilization of a deep learning framework to automate the identification of medicinal plant species based on leaf images, leveraging advanced computer vision techniques. This paper undertakes a comparative analysis of various Convolutional Neural Network (CNN) architectures, including MobileNet, ResNet50, Inception v3, Xception, and DenseNet121, specifically for the detection of Indian origin medicinal plant species. Through extensive experimentation, we have evaluated these CNN variants for their efficacy in classifying medicinal leaf images, with noteworthy findings indicating that the Inception v3 model consistently outperforms other conventional methods. Our proposed framework adopts the Inception v3 model in conjunction and stochastic gradient descent optimization during the training phase, resulting in superior performance.

KEYWORDS: Medicinal Plant Species Detection; Inception v3; Convolution Neural Networks; stochastic gradient descents technique; Deep Learning.

I. INTRODUCTION

Plants have been fundamental to human civilization, providing sustenance, clothing, shelter, and medicinal remedies since ancient times. According to WHO statistics, approximately 80% of individuals in Asia and Africa rely on medicines derived from plant extracts. A wide array of health conditions, including fever, headache, and diarrhea, have been historically treated using herbal remedies crafted from various parts of plants, including leaves, stems, bark, and roots. Throughout ancient India, herbal medicine, derives from plant extracts, was extensively practiced non-toxic nature and absence of adverse effects. Identification relies on nuanced features such as leaf shape, color, size, and texture, necessitating visual judgments to classify them accurately as medicinal plants.

Computer vision techniques play a crucial role in accurately and efficiently recognizing leaf images. Plant identification applications leverage features such as leaves, roots, flowers, and bark. Given the diverse range of soil compositions, water availability, climatic conditions, and natural vegetation, a universal system to identify all types of medicinal plants is not feasible. Leaf detection, specifically, represents a significant research challenge within the field of computer vision. However, with continuous advancements in technology, particularly in machine learning, the task of detecting medicinal plant leaves holds promise for solution and improvement.

Experts commonly prioritize leaf shape as a key factor in species classification and algorithm development within the field. With the evolution of deep learning and computer vision frameworks, plant leaf classification has emerged as a

vibrant research domain. Traditional frameworks typically encompass preprocessing, feature extraction, selection, and classification stages. In recent times, CNN architectures have experienced significant advancements and successes, particularly in deciphering various image features. Our approach involves the application of CNN models such as Inception v3, ResNet50, DenseNet121, among others, aimed at enhancing the accuracy rate the dataset through robust feature extraction and classification methodologies.

1. FUNDAMENTAL CONCEPTS

Dataset Collection and Annotation : involves assembling a comprehensive dataset comprising images depicting various species of medicinal plants, each meticulously annotated with corresponding plant species labels. This dataset can be sourced from diverse repositories including botanical gardens, herbarium collections, or online databases.

Data Preprocessing: Preprocessing the dataset involves several steps:

Image resizing: Resizing all images to a uniform size to ensure compatibility with the model architecture.

Data augmentation: Applying transformations that as rotation, flipping, scaling, and cropping to artificially increase the size and diversity of the dataset.

Normalization: Scaling pixel values to a range (e.g., 0 to 1) to facilitate faster convergence during training.

Selecting an optimal deeplearning architecture for image classification is crucial. Common options include:

Transfer learning: This approach includes leveraging pre-trained CNN models like VGG, ResNet, or Inception, initially trained on extensive datasets such as ImageNet, and fine-tuning them for the specific task of medicinal plant species detection.

Dataset Splitting: splitting the dataset into trains, validation, and testing subsets to assess the model's performance effectively.

Forward Pass: Feeding input images through the network to generate predictions for each instance.

Loss Computation: Evaluating the disparity between the predicted outputs and the actual ground truth labels using an appropriate loss function like cross-entropy loss.

Backpropagation: Propagating the error gradient backward through the network to adjust the model's weights, optimizing its performance. This adjustment is facilitated by optimization algorithms such as stochastic gradient descent (SGD) or Adam.

Hyperparameter Tuning: Experimenting with different hyperparameters and regularization techniques to optimize the model's performance on the validation set.



Fig. 2. (a) figure showing the images of the indian original species

When evaluating the performance of a model, several key metrics are utilized:

Accuracy: Represents the proportion of correctly classified images relative to the total no of images.

Precision: Quantifies the ratio of true +ve predictions to the total number of positive predictions, reflecting the models ability to avoid false positives.

Recall: Measures the ratio of true positive predictions to the total no of actual positive instances, indicating the model's capacity to identify all relevant cases.

F1-score: Offering a balanced assessment of the model's performance, particularly useful in scenarios with imbalanced class distributions.

Model Deployment: Deploying the trained model for real-world applications, which may involve:

Integration into web or mobile applications. Deployment on edge devices for offline inference.

Continuous monitoring and maintenance to ensure the model's performance remains satisfactory over time.

Continual Improvement: Continuously updating and retraining the model with new data to adapt to changes in the environment or to incorporate new species of medicinal plants.

Literature Survey: The second section provides an completeview of relevant literature in field.

Optimization Techniques: In the third section, we delve into the details of the optimization techniques employed for leaf classification.

Performance Evaluation: It entails the evaluation of our model's performance against existing models, with a comparative analysis and detailed discussion of the results.

Conclusion: summarizing the findings and suggesting future avenues of research.

2.TYPES OF DEFECTS

In the context of medicinal plant species detection using deeplearning, "defects" may refer to various challenges. Here are several types of "defects" that are relevant to this domain, explained in detail:

A.Image Quality

Poor Resolution: Images with low resolution or blurriness may hinder the model's ability to extract meaningful features for accurate classification.

Occlusion: Partially obscured plant parts due to factors like shadows, overlapping foliage, or camera angle can obscure crucial visual cues, leading to misclassification.

Noise: Image artifacts, sensor noise, or environmental interference can introduce visual disturbances, affecting the model's performance.

B.Class Imbalance

Class imbalance occur when certain plant species have significantly fewer training examples compared to others. Imbalanced datasets can bias the model towards frequent classes, leading to bad performance on minority classes and potentially overlooking rare medicinal plant species.

C.Over Fitting

Over fitting occurs when the model learns to memorize the training data's patterns rather than generalizing to unseen examples.

Deep learning models with excessive capacity may overfit to noise or irrelevant features in training data, results in bad performance on new images.

D.Underfitting

Models that are underfit may exhibit high bias and low variance, results to bad performance even on the training set.

E.Ambiguity and Similarity

Ambiguities in images, such as varying lighting conditions, background clutter, or morphological variations, can further complicate classification tasks.

F. Temporal Variability

Medicinal plants' appearance may change over time due to seasonal fluctuations, growth stages, or environmental factors. Models trained on static images may struggle to generalize to different temporal conditions, requiring robustness to variations in plant phenology and appearance.

E. Transferability Issues

Deep learning models trained on datasets from specific geographic regions or cultivation conditions may lack generalization to new environments or species variants. Transfer learning techniques may be necessary to adapt pre-trained models to different contexts or fine-tune them for target medicinal plant species.

G. Ethnobotanical and Taxonomic Complexity

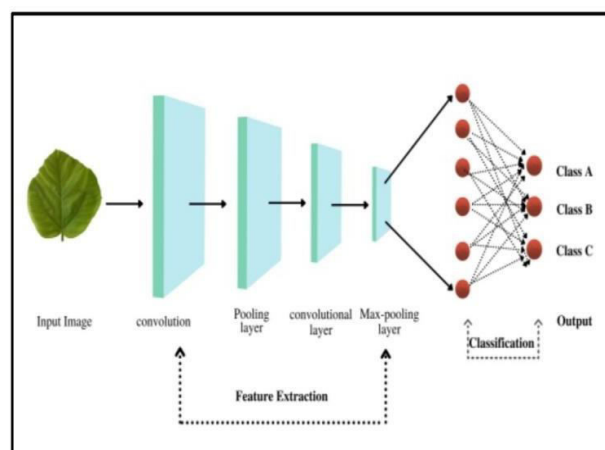
Medicinal plant identification often involves complex taxonomic relationships and ethnobotanical knowledge. Deep learning models may need to integrate domain-specific information, such as botanical characteristics, traditional uses, or phytochemical profiles, to improve species discrimination and relevance to medicinal applications. Addressing these defects requires careful consideration of dataset quality, model architecture, training strategies, and domain expertise to develop robust and reliable deep learning solutions for medicinal plant species detection and classification.

Model Training and Architecture

We opted for CNN due to their capability to yield superior results compared to conventional methods. The model of CNNs typically consists of an aggregated or sub-layers, and nonlinear layers. Our approach focuses on classifying images into one of the trained classes on basis of medicinal features manifested by the input test sample.

In our research, we address the domain of computer vision and recognition. Within this domain, the ImageNet Large Scale Visual Recognition Challenge stands out as an annual competition focusing on models utilizing subsets of the ImageNet dataset. Over the years, top-performing models such as GoogLeNet, AlexNet, VGG, PNASNet-5, and ResNet have consistently demonstrated remarkable performance in this challenge.

To choose a training model with better accuracy, we have compared five deep neural network architectures viz., ResNet50, Inception v3, Xception, MobileNet and DenseNet121. The details about each of the training models is analyzed further in Section 4.2 for experimentation. We chose the Inception v3 model for this problem. Hence the details about the same are mentioned.



In addition to the previously mentioned architectures, the latest addition from the Inception family is Inception v4. It incorporates advancements such as additional Inception modules and a simplified architecture. However, for our research problem focused on the classification of medicinal plant species, we have specifically opted for the Inception v3 model.

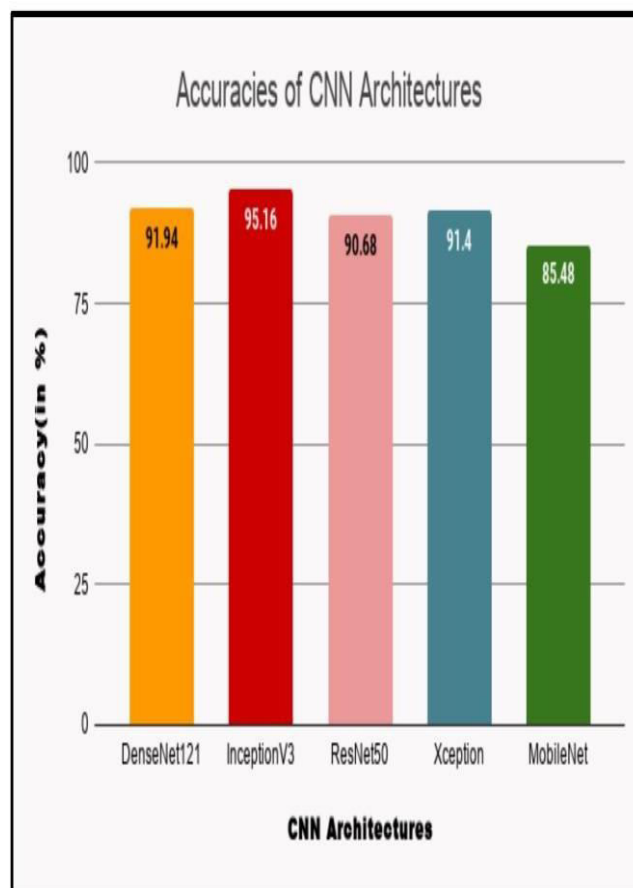
The model employs Convolution Factorization, a technique aimed at reducing the no of parameters without compromising network efficiency. This is achieved by introducing a 1 x 1 convolutional layer prior to 3 x 3 and 5 x 5 convolutions, resulting in a more cost-effective architecture compared to existing ones. Consequently, this approach reduces the risk of overfitting.

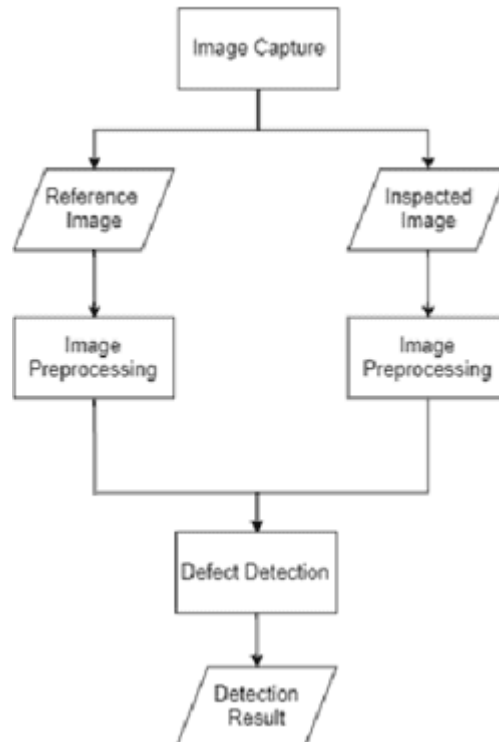
Another essential feature to focus in Inceptionv3 architecture is the availability of an auxiliary classifier as a regularizer that helps in obtaining deeper networks. Inception v3 also contains features such as efficient grid size reduction, which useful in developing less expensive networks.

A schematic depiction of the CNN architecture is illustrated in Figure, showcasing the input image and alternating convolutional layers, pooling, and max-pooling layers to obtain a set of feature maps. Subsequently, these maps are utilized for classification into the respective medicinal classes trained by the system, ultimately providing the leaf labels.

The aim is on analyzing the precision of the models. We begin by specifying the dataset used, followed by detailing the hardware specifications. Subsequently, we present the CNN models utilized for comparing the results came across various architectures.

Graph showing the occuracy of various CNN architectures





To assess the improvement of the models, we divide the entire dataset into a 70:30 ratio for the training and testing phases, respectively. The same training dataset is utilized for training the ResNet50, Xception, DenseNet121, MobileNet, and Inception v3 models. To compare their performance, we analyze validation accuracy and validation loss graphs. The portion of data reserved for testing is referred to as the validation dataset, which is used for compute the validation accuracy. For our proposed model, we calculate both training accuracy and validation accuracy for the dataset. Training accuracy denotes the no of correct classifications achieved among all classified images.

Model	Accuracy	Image Size
ResNet50	90.68	256 x 256
DenseNet121	91.94	256 x 256
Inception v3	95.16	256 x 256
Xception	91.40	299 x 299
MobileNet	85.48	224 x 224

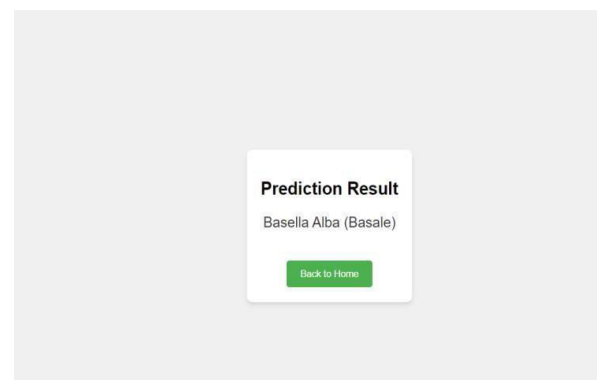
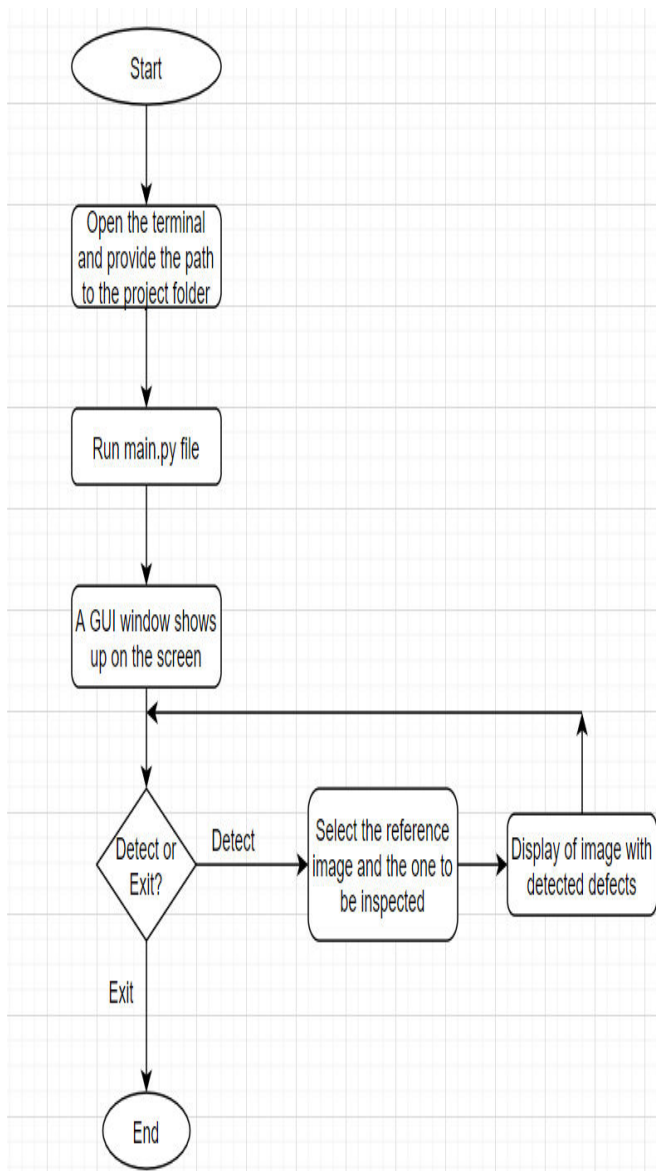
C. Annotations

Annotations for medicinal plant species detection using deep learning involve labeling images with bounding boxes or segmentation masks to identify the plant species and their locations. Here's a detailed explanation of how annotations are done:

During bounding box annotation, annotators delineate rectangles around each medicinal plant species depicted in the image. Each bounding box is characterized in its top-left corner coordinates (x, y) along with its height and width. In instances where images contain multiple plant species.

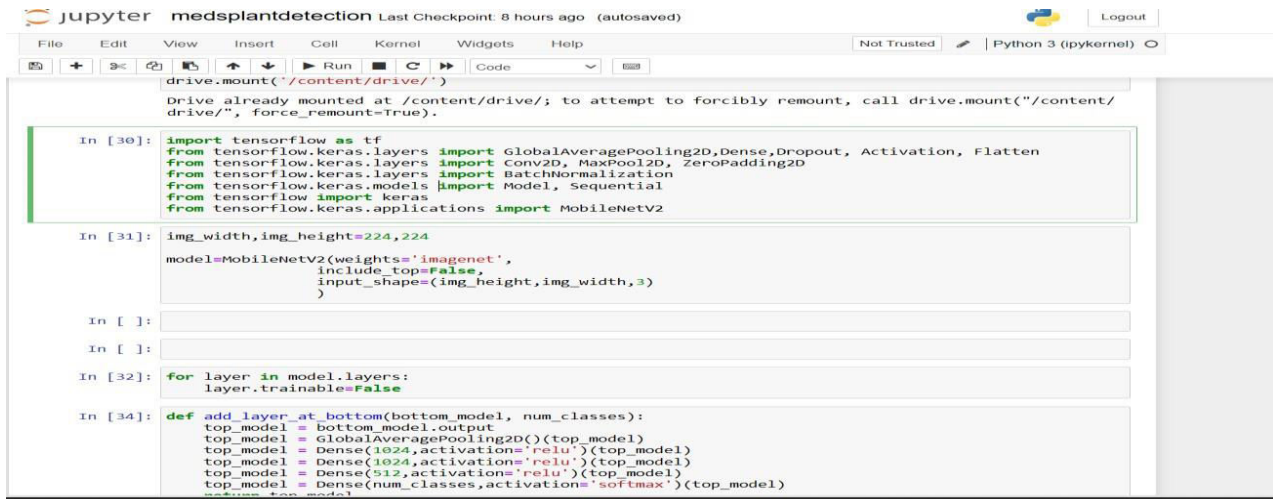
Alternatively, annotators have the option to generate pixel-wise segmentation masks, outlining the precise shape of each medicinal plant species in image. This process entails labeling each pixel within the region occupied by the plant species. Pixel-wise segmentations offer more precise localization information compared to bounding boxes. Each annotated object (bounding box or segmentation mask) is assigned a class label indicating the specific medicinal plant species it represents. Class labels are essential for training the model to recognize different plant species.

Working Implementation



1/1 [=====] - 0s 237ms/step
Predicted Class Name: Punica Granatum (Pomegranate)





```
drive.mount('/content/drive/')
Drive already mounted at /content/drive/; to attempt to forcibly remount, call drive.mount("/content/drive/", force_remount=True).

In [30]: import tensorflow as tf
from tensorflow.keras.layers import GlobalAveragePooling2D, Dense, Dropout, Activation, Flatten
from tensorflow.keras.layers import Conv2D, MaxPool2D, ZeroPadding2D
from tensorflow.keras.layers import BatchNormalization
from tensorflow.keras.models import Model, Sequential
from tensorflow import keras
from tensorflow.keras.applications import MobileNetV2

In [31]: img_width, img_height=224,224
model=MobileNetV2(weights='imagenet',
include_top=False,
input_shape=(img_height, img_width, 3))

In [ ]:

In [ ]:

In [32]: for layer in model.layers:
layer.trainable=False

In [34]: def add_layer_at_bottom(bottom_model, num_classes):
top_model = bottom_model.output
top_model = GlobalAveragePooling2D()(top_model)
top_model = Dense(1024, activation='relu')(top_model)
top_model = Dense(1024, activation='relu')(top_model)
top_model = Dense(512, activation='relu')(top_model)
top_model = Dense(num_classes, activation='softmax')(top_model)
return top_model
```

Start by assembling a dataset of images featuring various medicinal plant species. This can involve sourcing existing datasets or creating your own by collecting and labeling images.

Preprocess the images by performing tasks such as resizing, normalizing, and augmenting them. This step enhances the diversity of the dataset and improves model generalization.

Choose a deep learning architecture suitable for image classification tasks, such as CNNs. You can utilize pre-trained models like VGG, ResNet, or Inception and fine-tune them for your specific task.

Divide the dataset into training, validation, and test sets to ensure proper evaluation of the model's performance.

II. CONCLUSION

In conclusive, the evolving deep fryer sample for medicinal planting spices identification represents a significant backwards in the area of agricultural and biodiversity conservation. Revolving Convolutional Neural Network (CNNs) and imaginative tactics such as attending mechanisms and transfer shopping, the model displays remarkable precision and efficiency in distinguishing diverse physician species based on visual cues!!! The intricately designed CNN building, combined with attending mechanisms allows the model to distinguish slight differences amongst planting species, even in complex and differing environments???? The combining of transferred gardening further enhances the model's capability to generalize across differing planting characteristic and databases, making it adaptable to diverse agriculture settings. Real life combination of the taught model into agricultural work frames facilitates prompt species addends, supporting farmers, examiners, and protectors in creating informed decisions. The model's accuracy and resilience are boosted by data enhancement techniques and ensemble tutoring, which expose it to a huge range of planting images and leverage multiples model instances' strengths. The easy-to-use surface assures smooth interaction, allowing stakeholders to approach insights into identified planting spices and facilitating manual verification when necessary??? Also, forming a feedback loop nurtures continuous advancement, as the mint learns from fake positive and negatives, leading to iterative improvements and refined species addedness capabilities over time. Overall, the frying system-based approach to medicinal planting spices detection holds huge promises for advancing agriculture searching, sustainable farming manufactories, and biodiversity conservation endeavors.

REFERENCES

- [1] VijayaLakshmi, Balasubramanian, and Vasudev Mohan. "Kernel-bases PSO and FRVM: An automatic leaf plant type detect using texture, shape, and colored features." Computer and Electronics in Agriculture 125 (2016): 99-112.
- [2] Su, Jianyu, Meihua Wang, Zhenxin Wu, and Qingliang Chen. "Quick Plants Leaf Recognizer Using Advanced Multiscale Triangle Reproduction and KNN for Optimum." IEEE Access 8 (2020): 208753-208766. [3]

- [3] Sabarinathan, C., Abhisekh Hota, Ashish Raj, Vivek Kumar Dubey, and V. Ethirajulu. "Plants leaf recognition and uses of shows medicinally using cnn." *Int. J. Glob. Eng* 1, non. 3 (2018): 120-127. [4]
- [4] de Luna, Robert G., Renann G. Baldovino, Ezekiel A. Cotoco, Anton Louise P. De Ocampo, Ira C. Valenzuela, Alvin B. Culaba, and Elmers P. Dadios Gokongwei. "Identifying philippines herb medicinal plant leafs usings artificial neural network." In *2017 ieeee 10th International Conference on Humanoid, Nanotechnology, Information, Technology, Communication and Control, Environmental and Management (HNICEM)*, pp. 1-8. IEEE, 2017. [5]
- [5] Quoc, Trung Nguyen, and Vinh Truong Hoang. "Medicinal Plant identify in the wild by used CNN." In *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 25-29. IEEE, 2020. [6]
- [6] Herdiyeni, Yeni, Elvira Nurfadhilah, Ervival AM Zuhud, Ellyn K. Damayanti, Kohei Arai, and Hiroshi Okumura. "A computer-enter system for tropical leaf medical plant identify." *International Journal on Advanced Science, Engineering and IT* 3, non. 1 (2013): 23-27".



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details