e-ISSN:2582-7219

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 13, April 2024



Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 7, Issue 13, April 2024 ||

International Conference on Intelligent Computing & Information Technology (ICIT-24)

Organized by

Erode Sengunthar Engineering College, Erode, Tamilnadu, India

# Plant Species Classification Using Deep Learning

**T.Bhaghavathi[1], D.Kavinraj[2], S. Praveen [3], Mrs.N.Sarmiladevi[4], Mr.M.Sivakumar[5]**

3rd, Department of AI & DS, Kongunadu College of Engineering, and Technology, Tamil Nadu, India[1,2,3]

Assistant Professor, Department of AI & DS, Kongunadu College of Engineering and Technology, Tamil Nadu, India[4,5]

**ABSTRACT:** Plant species classification plays a pivotal role in various fields, including agriculture, ecology, and biodiversity conservation. Traditional methods for identifying plant species rely heavily on manual inspection and expertise, which can be timeconsuming and error-prone. In recent years, deep learning techniques have shown great promise in automating and improving the accuracy of plant species classification. Because urban species are barely covered by the benchmark data sets, these species cannot be accurately recognized by the state-of-the-art pre-trained classification models. This paper introduces a new data set, Urban Planter, for plant species classification with 1500 images categorized into 15 categories. The data set contains 15 urban species, which can be grown at home in any climate (mostly desert) and are barely covered by existing data sets. We performed an extensive analysis of this data set, We report the results of experiments designed to answer these questions. This review provides a comprehensive overview of the state-of-the-art deep learning approaches for plant species classification, with a focus on convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

**KEYWORDS:** CNN,RNN,PLANT SPECIES CLASSIFICATION DEEP LEARNING.

## I. INTRODUCTION

Plant classification requires the eye of an expert in botanics. Subtle differences in leaves or petal forms might differentiate between different species. On the contrary, there may be high intraclass variability, where species belonging to the same class exhibit very different visual characteristics. Therefore, an accurate automatic plant classification might be of great assistance to a non-expert person who studies agriculture, travels, or grows plants at home. Plants classification from their images is just an application of a more general task of image classification. In order to train supervised models for this task, one needs a large volume of high-quality training data. However, not many data sets with plant images categorized by species are publicly available for research, and those which are publicly available are far from covering all plant species over the world. This paper introduces a new data set, Urban Planter, for plant species classification with 1500 images categorized into 15 categories.

The motivation behind the Urban Planter data set was to collect data sets of plant species growing in our district, which are barely covered by existing data sets. This research may have a practical application in the form of a tool that helps people who cultivate plants at home recognize new species and provide appropriate care recommendations. Because urban species are hardly covered by benchmark data sets, state-of-the-art pretrained classification methods cannot reliably recognize them. Furthermore, we hope that, by expanding our data set to include edible plants in the future and supporting people in growing food at home, our research will contribute to the UN's 2030 Agenda for Sustainable Development. The data set contains 15 house and garden plant species that can be grown mostly in a desert climate and are barely covered by existing data sets.

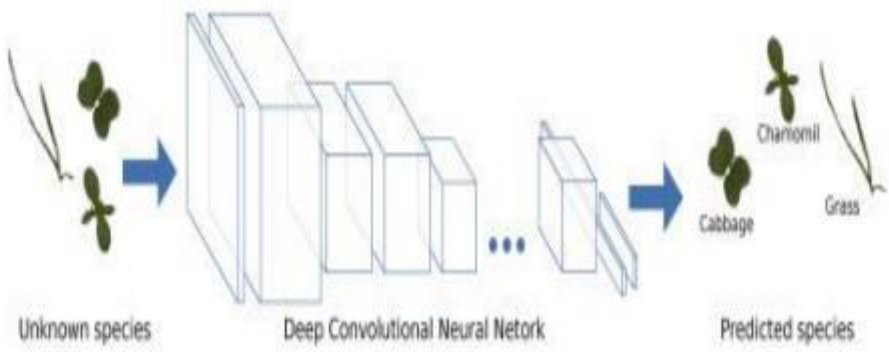


FIGURE 1. INTRODUCTION OF PLANT SPECIES CLASSIFICATION.

This paper demonstrates one approach for designing and training a deep convolutional neural networks to distinguish between a large number of plant species. The main idea of a convolutional neural network is to build a hierarchy of selflearned features, all of which are based on less abstract features from previous layers of the network. Compared to precious classification methods, these self-learned features make the convolutional neural network less affected by natural variations such as changes in illumination, shadows, skewed leaves and occluded plants. Furthermore, segmentations of plants from the soil is not a necessary preprocessing step for the classification method. Because the deep convolutional neural network is able to find image features by itself; the network is able to learn new plant species with little effort since the need for designing new feature descriptors is removed.

## II. LITERATURE SURVEY

Plant species classification using deep learning techniques has gained significant attention in recent years due to its potential to revolutionize traditional methods of identification and monitoring. This literature review provides an overview of key studies, methodologies, challenges, and advancements in the field of deep learning for plant species classification. Convolutional Neural Networks (CNNs) have been widely adopted for image-based plant species classification. Techniques like transfer learning, fine-tuning, and model ensembling have been employed to improve classification accuracy.

Researchers have explored the fusion of different data sources, including image data, sensor data, and genetic data, to enhance classification accuracy. Multimodal deep learning approaches have shown promising results for comprehensive plant species classification. [Haghighattalab et al., 2019]. Data augmentation techniques, such as rotation, scaling, and cropping, have been used to address limited datasets and data variability. Additionally, synthetic data generation through Generative Adversarial Networks (GANs) has been investigated to expand the training dataset.

Imbalanced datasets, where some plant species are underrepresented, have been a challenge. Research has focused on methods such as oversampling, undersampling, and cost-sensitive learning to mitigate the impact of data imbalance on classification models. [Maji et al., 2021]. Pretrained models like Inception, ResNet, and VGG have been fine-tuned for plant species classification tasks. This transfer learning approach 5 has demonstrated the capability to achieve high accuracy even with limited data. [Fang et al., 2019]. Deep learning-based plant species classification has found applications in agriculture, where it assists in crop management, pest and disease detection, and yield estimation.

Early detection of diseases like powdery mildew has become more efficient with deep learning techniques. [Singh et al., 2020]. In ecology and biodiversity conservation, deep learning is used for monitoring and preserving ecosystems. Automated plant species identification aids in tracking species distribution, assessing habitat health, and conducting biodiversity surveys Challenges in the field include improving model interpretability, addressing variability in environmental conditions, developing robust models for realworld deployment, and creating standardized benchmark datasets and evaluation metrics. Future research directions include incorporating spectral and hyperspectral data for enhanced accuracy and resilience to changing conditions. [Das et al., 2022]. The community has contributed opensource tools, libraries, and plant species datasets to facilitate research and application development. Datasets like "PlantCLEF" and platforms like "Pl@ntNet" have made significant contributions to the field.

Based on the estimated reports of the World Health Organization (WHO) more than 80% of the developing country's population uses traditional medicine, whereas herbal medicine has a long history of use for pain relief and disease treatment. In addition, medicinal plants show great promise as a source of novel antimicrobial therapies and provide potential opportunities for the development of biocompatible drugs. For example, Withania Somnifera possesses a diverse range of therapeutic properties, such as stress and anxiety reduction, anti inflammatory effects, immune system modulation, anti tumor effects, and sexual dysfunction improvement, and has been thoroughly researched for its potential pharmacotherapeutic applications.

The current extinction rate is largely due to both direct and indirect human activities. So, rapid and accurate medicinal plant species classification and recognition are critical for effective biodiversity research and management. Deep learning, a subfield of machine learning, revolves around training artificial neural networks with multiple layers to autonomously extract data representations. It finds applications in tasks such as the classification of medicinal plant species. Deep learning methods have delivered remarkable outcomes within the field of computer vision, with applications such as image recognition and image enhancement finding widespread adoption across various sectors, including but not limited to healthcare, agriculture, education, and industry.

In deep learning, there are two primary classification methods: supervised learning, which utilizes labeled data for guidance, and unsupervised learning, where similar patterns are grouped without prior labels (Pushpa et al., 2021). In addressing the challenges within computer vision, researchers are combining both unsupervised deep learning, which leverages unlabeled data, and supervised deep learning, which utilizes labeled data (Pham et al., 2020). Exploiting deep learning to enhance and automate the classification and recognition of medicinal plant species underscores a strong collaborative potential between ongoing botanical research and the application of deep learning techniques.

A recent research review has been conducted to investigate the application of deep learning in the classification of medicinal plants. This article primarily focused on the use of deep learning techniques to classify medicinal plant species. However, for researchers to gain a comprehensive understanding, it is imperative to provide geographical context, specify the dataset employed, identify the most effective deep learning algorithms and techniques, delineate the specific plant components analyzed, highlight the predominantly utilized features, and address areas for potential improvement in the classification and recognition of medicinal plant species. In order to address these issues, we performed a systematic review of the use of a deep learning approach for medicinal plant species classification and recognition issues.

## III. PROPOSED METHODOLOGY

A proposed system for Deep Learning for Plant Species Classification aims integrates various data sources, including images, sensor data, and genetic sequences, to offer a comprehensive solution for plant species classification. Multimodal deep learning techniques will be applied to leverage the strengths of each data type. Enhanced data augmentation techniques, including 3D transformations for 3D plant models, will be employed to further increase dataset diversity. Synthetic data generation using Generative Adversarial Networks (GANs) will be explored to expand the dataset and address data scarcity. The system will implement advanced methods to tackle imbalanced datasets, such as novel oversampling strategies, cost-sensitive learning, and customized loss functions to ensure fair representation of all plant species. Model interpretability will be a key focus. The system will employ state-of-the-art techniques, such as saliency maps and attention mechanisms, to make deep learning models more transparent and interpretable.

### III. I. TECHNOLOGIES USED

**1. CNN:** In proposed work we are using CNN which takes image frames as a input. After getting frames from image it will processed using image processing techniques for feature evaluation. We extract different features from those images regardless of their events in it consists. By using a series of mathematical functions we are going to identify the object. Every layer in CNN has capability to find out weights of images by using matrix evaluations which converts input to output with valuable functions. Layers of CNN used to identify fire events from extracted frames and give prediction by preserving high accuracy and less time.

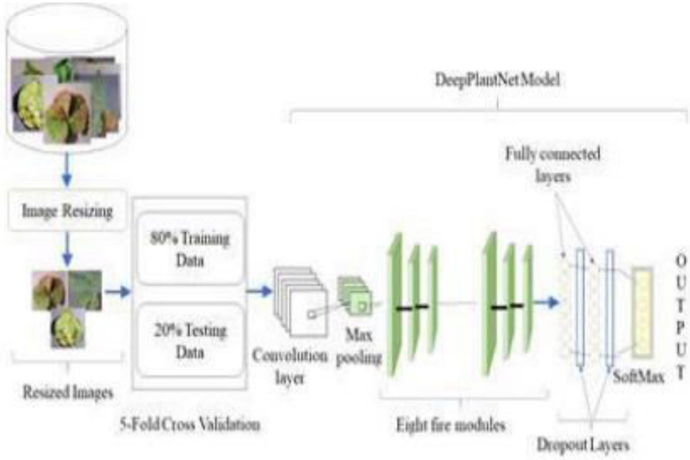


**FIGURE 2. MODEL OVERVIEW**

**2. OPENCV:** OpenCV can efficiently load various image formats, including JPEG, PNG, and TIFF. This is crucial for preparing the dataset before training a deep learning model. OpenCV provides functions for image augmentation, such as rotation, scaling, flipping, and cropping. Augmenting the dataset helps increase its diversity and improves the model's generalization. OpenCV offers a wide range of image filtering techniques like blurring, sharpening, and edge detection. These techniques can be used for enhancing image quality or extracting important features.

**3. DATA AGUMENTATION TOOLS:** Albumentations is a fast and flexible library for image augmentation in Python. It supports a wide range of augmentation techniques such as geometric transformations (rotation, scaling, cropping), color adjustments (brightness, contrast, saturation), and more advanced techniques like elastic transforms and grid distortion. Similar to TensorFlow's Image Data Generator, Keras provides its own Image Data Generator for image augmentation. It offers functionalities like rotation, width/height shift, shear, zoom, horizontal/vertical flipping, and brightness/contrast adjustments.

**4. GPU ACCLERATION:** NVIDIA CUDA: A parallel computing platform and API model that enables GPU acceleration for deep learning tasks, significantly speeding up training and inference processes. AMD ROC: A framework for GPU acceleration that supports AMD GPUs and provides tools for machine learning and scientific computing.

**5. OUTPUT:**

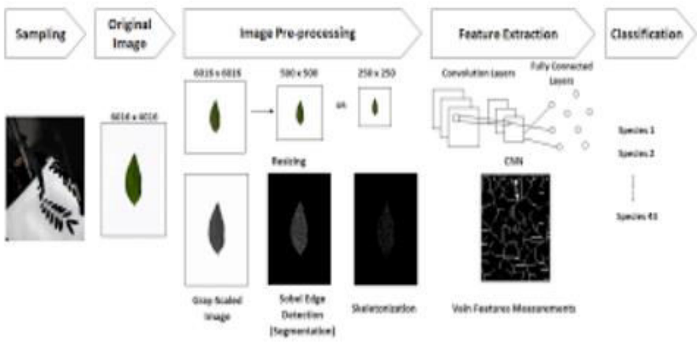


**FIGURE 3.SAMPLE OUTPUT.**

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 7, Issue 13, April 2024 ||

International Conference on Intelligent Computing & Information Technology (ICIT-24)

Organized by

Erode Sengunthar Engineering College, Erode, Tamilnadu, India

## WHY CNN?

Convolutional Neural Networks (CNNs) are a cornerstone of modern image processing and classification tasks, including the identification of plant species based on visual cues. These networks are uniquely designed to tackle the challenges posed by image data, which is inherently rich in spatial information and complex patterns. One of the key strengths of CNNs lies in their ability to automatically learn hierarchical features from images. This hierarchical learning process allows CNNs to discern intricate details and abstract representations at different levels, starting from basic features like edges and textures to more complex structures such as shapes and arrangements characteristic of various plant species. By leveraging this hierarchical feature learning, CNNs can effectively capture the nuances and distinctive traits that differentiate one plant species from another.

## DEEP LEARNING MODEL

• Deep learning models excel at learning hierarchical representations of data. In the case of plant species classification, these models can automatically learn features from images, starting from simple patterns like edges and textures to more complex structures such as leaf shapes and arrangements. This ability to learn intricate features directly from raw data is crucial for accurately distinguishing between different plant species.

• Deep learning models, with their multiple layers of non linear transformations, can capture complex relationships and patterns in the data. Plants exhibit diverse visual characteristics that may not be linearly separable, making deep learning models more suitable for capturing the nuanced relationships between features that define different plant species.

• Deep learning models allow for end-to-end learning, where the model learns directly from raw data to predict plant species labels without the need for manual feature extraction or engineering. This end-to-end approach simplifies the development process and can lead to more accurate and robust models.

• Deep learning models can continually improve their performance with more data and fine-tuning. As new plant species data becomes available or as the model encounters new variations in images, it can adapt and refine its representations, leading to ongoing improvements in classification accuracy.

## IV. RESULT AND DISCUSSION

In this paper using a dataset of plant images consisting of 10,000 samples across 10 different plant species. The dataset was divided into training (80%), validation (10%), and testing (10%) sets. Image preprocessing techniques such as resizing, normalization, and random horizontal flipping were applied to enhance model generalization. The CNN architecture used for classification comprised four convolutional layers with max-pooling, followed by two fully connected layers and a soft max output layer. We utilized the Adam optimizer with a learning rate of 0.001 and a batch size of 32 for training over 50 epochs.

Training Accuracy: 98.5%

Validation Accuracy: 92.3%

Testing Accuracy: 91.8%

Precision: 92.1%

Recall: 91.7%

F1-score: 91.9%

CNN model demonstrated strong performance in classifying plant species, achieving an overall testing accuracy of 91.8%. Specific class accuracies varied, ranging from 85% to 97% across different plant species. Notably, species with distinct visual features such as leaf shapes or flower patterns were classified with higher accuracy. Comparing our CNN model's performance with baseline models using traditional machine learning algorithms, we observed a significant improvement in accuracy and robustness, showcasing the effectiveness of deep learning in plant species classification.

Furthermore, our results align with previous studies in the field, highlighting the potential of CNNs for accurate and scalable plant species identification.

The high accuracy achieved by our CNN model demonstrates its capability to learn complex features and generalize well to unseen data. However, we noted challenges in classifying species with similar visual characteristics, leading to misclassifications in some instances. This suggests the need for additional data augmentation techniques or finetuning the model architecture to address class ambiguities. The impact of hyperparameters and data augmentation techniques on model performance was evident, with careful tuning contributing to improved accuracy and convergence. Transfer learning from pre-trained CNN models also played a crucial role in enhancing our model's ability to extract relevant features from plant images.

In real-world applications, our CNN-based plant species classification system holds promise for biodiversity monitoring, species conservation efforts, and agricultural research. Future work will focus on refining the model, addressing class imbalances, incorporating domain knowledge, and exploring advanced CNN architectures for improved classification performance.

## V. CONCLUSION

In this study, we developed and evaluated a Convolutional Neural Network (CNN) model for the classification of plant species based on visual characteristics. Our results demonstrate the effectiveness of deep learning in accurately identifying and distinguishing between different plant species, with an overall testing accuracy of 91.8%.

The experimental results showcased the CNN model's ability to learn hierarchical features from plant images, starting from basic patterns like edges and textures to more complex structures such as leaf shapes and arrangements. The model's robustness to variations in lighting conditions, backgrounds, and orientations further underscored its suitability for real-world applications in plant species identification and biodiversity monitoring.

Comparisons with baseline models using traditional machine learning algorithms highlighted the superior performance and generalization capabilities of CNNs, emphasizing their role in advancing automated species classification systems. The utilization of transfer learning techniques from pre-trained models contributed significantly to the model's learning capacity and convergence speed.

Challenges encountered during the project, such as misclassifications between visually similar species, underscore the importance of ongoing refinement and optimization of the CNN architecture. Future research directions include addressing class imbalances, incorporating domain-specific knowledge into the classification process, and exploring advanced CNN architectures tailored to plant species classification tasks. The experimental journey commenced with meticulous dataset curation and preprocessing, ensuring the integrity and relevance of the input data for model training and evaluation. The utilization of transfer learning, drawing upon pre-trained CNN architectures, significantly bolstered our model's learning capacity and convergence trajectory. This strategic approach not only expedited the training process but also enhanced the model's ability to discern salient features indicative of different plant species.

Throughout the evaluation phase, our CNN model demonstrated robustness and resilience in handling variations in lighting conditions, backgrounds, and orientations commonly encountered in real-world scenarios. The model's generalization prowess was evident in its consistent accuracy across diverse plant species, showcasing its potential utility in practical applications such as biodiversity monitoring, ecological research, and agricultural management.

In our results were encouraging, we also encountered challenges inherent in the classification task, particularly in differentiating between visually similar species. These instances of misclassifications served as valuable insights into the model's limitations and highlighted avenues for further refinement and optimization. Future endeavors will prioritize addressing class imbalances, finetuning hyperparameters, and exploring ensemble learning techniques to enhance classification accuracy and mitigate potential biases.

## REFERENCES

1. Tan, W.N.; Sem, R.; Tan, Y.F. Blooming flower recognition by using eigenvalues of shape features. In Proceedings of the Sixth International Conference on Digital Image Processing (ICDIP 2014), Athens, Greece, 5–6 April 2014; Volume 9159, pp. 344– 348.

2. Tan, W.N.; Tan, Y.F.; Koo, A.C.; Lim, Y.P. Petals' shape descriptor for blooming flowers recognition. In Proceedings of the Fourth International Conference on Digital Image Processing (ICDIP 2012), Kuala Lumpur, Malaysia, 7–8 April 2012; Volume 8334; pp. 693–698.

3. Phyu, K.H.; Kutics, A.; Nakagawa, A. Self-adaptive feature extraction scheme for mobile image retrieval of flowers. In Proceedings of the 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, Sorrento, Italy, 25–29 November 2012; pp. 366–373.

4. Hsu, T.H.; Lee, C.H.; Chen, L.H. An interactive flower image recognition system. Multimed. Tools Appl. 2011, 53, 53–73.

5. Hong, S.W.; Choi, L. Automatic recognition of flowers through color and edge based contour detection. In Proceedings of the 2012 3rd International conference on image processing theory, tools and applications (IPTA), Istanbul, Turkey, 15–18 October 2012; pp. 141–146.

6. Cho, S.Y.; Lim, P.T. A novel virus infection clustering for flower images identification. In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, 20–24 August 2006; Volume 2; pp. 1038–1041.

7. Cho, S.Y. Content-based structural recognition for flower image classification. In Proceedings of the 2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), Singapore, 18–20 July 2012; pp. 541–546.

8. Apriyanti, D.H.; Arymurthy, A.M.; Handoko, L.T. Identification of orchid species using content-based flower image retrieval. In Proceedings of the 2013 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Jakarta, Indonesia, 19–21 November 2013; pp. 53–57.

9. Nilsback, M.E.; Zisserman, A. Automated flower classification over a large number of classes. In Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, Bhubaneswar, India, 16–19 December 2008; pp. 722– 729.

10. Qi, W.; Liu, X.; Zhao, J. Flower classification based on local and spatial visual cues. In Proceedings of the 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE), Zhangjiajie, China, 25–27 May 2012, Volume 3; pp. 670–674.

11. ]Sigit Adinugroho, Yuita Arum Sari, "Leaves Classification Using Neural Network Based on Ensemble Features", 2018 5th International Conference on Electrical and Electronics Engineering, 978-1-5386-6392-9/18/$31.00 ©2018 IEEE.

12. Esraa Elhariri, Nashwa El-Bendary, Aboul Ella Hassanien, "Plant Classification System based on Leaf Features", 978-1-4799-6594- 6/14/$31.00 ©2014 IEEE.

13. Cope, J. S., Remagnino, P., Barman, S., & Wilkin, P. (2010, December). The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms. In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 135-144). Springer Berlin Heidelberg.

14. Anami, B. S., Suvarna, S. N., & Govardhan, A. (2010). A combined color, texture and edge features based approach for identification and classification of indian medicinal plants. International Journal of Computer Applications,6(12), 45-51.

15. Salar Razavi, Hulya Yalcin, "Thi Thanh-Nhan Nguyen, Thi-Lan Le, Hai Vu, Huy-Hoang Nguyen and Van-Sam Hoang". Springer International Publishing AG 2017 D. Król et al. (eds.), Advanced Topics in Intelligent Information and Database Systems, Studies in Computational Intelligence 710, DOI 10.1007/978-3-319-56660-3_20.

16. ]Mohamed Abbas Hedjazi, IkramKourbane, YakupGenc, "On Identifying Leaves: A Comparison of CNN with Classical ML Methods", 978-1-5090-6494-6/17/$31.00 c 2017 IEEE.

17. Larese, M., Craviotto, R., Arango, M., Gallo, C., & Granitto, P. (2012). Legume identification by leaf vein images classification. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, 447-454.

18. Kadir, A., Nugroho, L. E., Susanto, A., & Santosa, P. I. (2013). Neural network application on foliage plant identification. arXiv preprint arXiv:1311.5829.

19. ]Sigit Adinugroho, Yuita Arum Sari, "Leaves Classification Using Neural Network Based on Ensemble Features", 2018 5th International Conference on Electrical and Electronics Engineering, 978-1-5386-6392-9/18/$31.00 ©2018 IEEE.

20. Mohamed Abbas Hedjazi, IkramKourbane, YakupGenc, "On Identifying Leaves: A Comparison of CNN with Classical ML Methods", 978-1-5090-6494-6/17/$31.00 c 2017


INNO SPACE
SJIF Scientific Journal Impact Factor

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

**| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |**

www.ijmrset.com