International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 02 Feb 2024, pp: 399-403 www.ijaem.net ISSN: 2395-5252

A Comprehensive Approach to Plant Identification: Innovations in Image-Based Taxonomy

Naiyani Paladugu[1], Ragi Yashaswini[2], Dr. K. Rajitha[3]
[1,2]*Student, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana.*
[3]*Assistant Professor,Mahatma Gandhi Institute of Technology, Hyderabad, Telangana*
Corresponding Author: Naiyani Paladugu

Date of Submission: 18-02-2024 Date of Acceptance: 29-02-2024

ABSTRACT: Plants serve as a crucial source of sustenance for a multitude of organisms, including humans, birds, animals, and insects. With the world's population on the rise and natural resources dwindling, there's a pressing need to enhance agricultural efficiency. Fortunately, the integration of modern computing technologies offers promising solutions across various agricultural domains. Among these advancements, leveraging Classification algorithms stands out as a potent tool for automating the identification of plant species, a process that is traditionally laborious and time-consuming. Convolutional Neural Networks (CNNs), a cornerstone of image processing, excel in discerning relevant features and patterns from visual data, enabling tasks like image classification, object detection, and segmentation. This advancement holds immense promise for biologists, students, environmentalists, and other stakeholders by streamlining plant identification processes and fostering deeper insights into plant biodiversity and conservation efforts.

KEYWORDS: Convolutional neural network, image classification

I. INTRODUCTION

With the global population steadily increasing and natural resources becoming scarcer, the agricultural sector faces mounting pressure to enhance efficiency and productivity. In this context, leveraging modern computing technologies presents a promising solution to address the challenges of agricultural sustainability and food security. Among these technologies, Convolutional Neural Networks (CNNs) have emerged as powerful tools for automated image recognition and classification tasks.

This project focuses on utilizing CNNs for the purpose of plant identification—an essential component of agricultural research, biodiversity conservation, and ecosystem management. The ability to accurately and efficiently identify plant species from images is paramount for various stakeholders, including biologists, farmers, environmentalists, and policymakers. Traditional methods of plant identification often rely on manual observation and expertise, which can be time-consuming, labor-intensive, and prone to errors.

By employing CNNs, this project aims to automate the plant identification process, thereby streamlining workflows and facilitating faster and more accurate species recognition. CNNs are well-suited for this task due to their ability to extract intricate features from images and learn complex patterns, enabling them to distinguish between different plant species with high precision.

Through this project, we seek to demonstrate the feasibility and effectiveness of CNN-based plant identification systems, showcasing their potential to revolutionize agricultural practices and contribute to broader efforts in biodiversity conservation and environmental stewardship. Additionally, we aim to provide insights into the underlying principles of CNNs and their application in real-world scenarios, fostering a deeper understanding of the intersection between artificial intelligence and agriculture.

II. PROBLEM DEFINITION

There are many plant species available to man, even some that have not been discovered. There are species of plants are reported to be on the verge of extinction. Due to the inhumane nature of humans, forest is cleared to make space for

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 02 Feb 2024, pp: 399-403 www.ijaem.net ISSN: 2395-5252

industrial purposes. These purposes will lead to the destruction of nature, including plants. Plants are important to be preserved for the future. Plants should be made aware to the public to avoid such disaster. Ensuring they know the plants species are the reasons for the development of this system. The system will ensure people can identify plants without having flipping through research papers or books. Those plants that are similar to one another can be recognized easily. Using convolutional neural network to recognize plant can help the public gain more awareness towards plants species. Using the traditional convolutional neural network for the training of model to recognize plants. The system was tested to get the best accuracy to recognize plants. The images were increased to ensure a better accuracy. This system requires a good model in order to recognize the correct plants. Convolutional neural network is commonly used for image classification due to its high accuracy. The accuracy for the system created in this project is 94%. This project can be improved in the future by increasing the accuracy. There are a lot of ways to increase the accuracy such as trying out other CNN architectures or by feeding the system more images.

III. EXISTING SYSTEM

This system helps users to identify plants from photographs of their leaves. The recognition engine consists of a backend server that accepts input images from various front-end clients. The data exchange between the clients and web service is ensured via HTTP protocol. Web service consists of a library called leaf core that implements leaf identification logic and a file system that stores leaf data set. The leaf core logic web service then processes the input image and extracts its relevant features. Identification is done by running a weighted K nearest neighbor's search on database images.

IV. PROPOSED SYSTEM

Convolutional neural network (CNN) is one of the main categories to do images recognition, images classifications, etc., If once the user opens the web application and upload the image. The CNN classifier process it and classify it under certain categories (E.g., Rice, Sugarcane, Maize, Wheat). Classifier sees an input image as an array of pixels and it depends on the image resolution. Based on the image resolution, it will see h x w x d (h = Height, w = Width, d = Dimension) array of the matrix of RGB and an image of 4 x 4 x 1 array of a matrix of the grayscale image. Each input image will pass through a series of convolution layers with filters (Kernal's), Pooling, fully connected layers (FC) and apply SoftMax function to classify an object with probabilistic values between 0 & 1. The above flowchart (fig. 1) is a complete flow of CNN to process an input image and classifies the objects based on values.

V. HARDWARE REQUIREMENTS

1. **Processor**: Intel Pentium® Dual Core Processor (Minimum)

A dual-core processor from Intel, such as the Pentium® series, is the minimum requirement for efficient processing of tasks related to the system's operations. This processor configuration ensures that the system can handle computational tasks effectively without significant performance bottlenecks.

2**. Speed**: 2.9 GHz (Minimum)

The minimum processor speed requirement of 2.9 GHz ensures that the system can execute tasks swiftly and efficiently. A higher clock speed enhances the overall responsiveness and performance of the system, allowing for smoother operation and reduced processing times for complex tasks

3**. RAM**: 2 GB (Minimum)

With a minimum of 2 GB of RAM (Random Access Memory), the system can store and access data efficiently during its operation. Adequate RAM capacity ensures smooth multitasking and facilitates the seamless execution of various processes, including image processing and classification tasks, without experiencing significant slowdowns or memory-related issues.

VI. SOFTWARE REQUIREMENTS

1.**Operating System**: Windows 7 (Minimum)

The system is compatible with Windows 7 as the minimum operating system requirement. Windows 7 provides a stable and reliable platform for running the system's software components and ensures compatibility with other essential software tools and drivers required for optimal performance

2**. Front End**: Streamlit

Streamlit serves as the front end for the system, providing a user-friendly interface for interacting with the plant identification application. Streamlit offers intuitive features and interactive elements that facilitate seamless navigation and exploration of plant images and classification results, enhancing the overall user experience

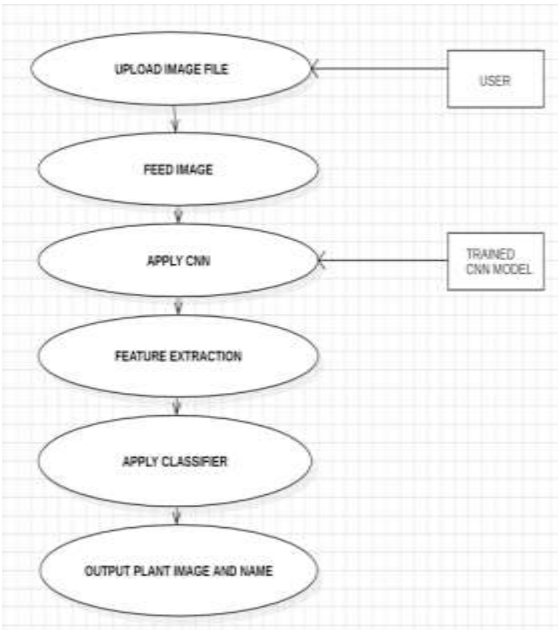
International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 02 Feb 2024, pp: 399-403 www.ijaem.net ISSN: 2395-5252

3. **Back End**: Python

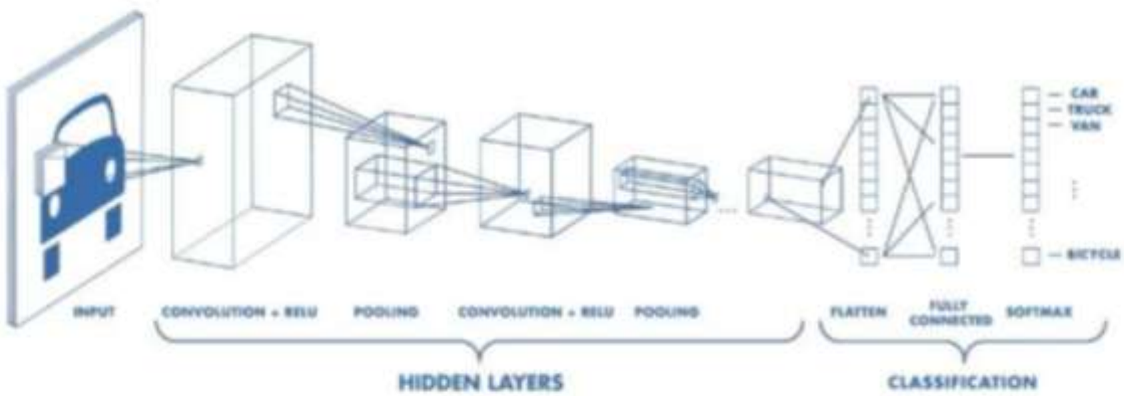
Python serves as the back end for the system, handling the computational tasks related to image processing, classification, and system functionality. Python's versatility and extensive libraries, such as TensorFlow and scikit-learn, enable efficient implementation of machine learning algorithms for plant identification. Additionally, Python's robust ecosystem supports integration with Streamlit for building interactive web applications with ease.

VII. SYSTEM ARCHITECTURE

Step1:
The first step is to collect data. We are using the "new dataset" Dataset.
Step2:
Pre-processing and Augmentation of the collected dataset is done.
Step3:
Building CNN (Convolutional Neural Network) Model for identification of various plants



VIII. CONVOLUTIONAL NEURAL NETWORKARCHITECTURE



A Convolutional Neural Network, also known as CNN or ConvNet, is a class of neural networks that specializes in processing data that has a grid-like topology, such as an image. A digital image is a binary representation of visual data. It contains a series of pixels arranged in a grid-like fashion that contains pixel values to denote how bright and what color each pixel should be.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 02 Feb 2024, pp: 399-403 www.ijaem.net ISSN: 2395-5252

A CNN typically has three layers: a convolutional layer, a pooling layer, and a fully connected layer.

Convolution Layer:

It carries the main portion of the network's computational load.This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.

During the forward pass, the kernel slides across the height and width of the image-producing the image representation of that receptive region. This produces a two-dimensional representation of the image known as an activation map that gives the response of the kernel at each spatial position of the image. The sliding size of the kernel is called a stride. If we have an input of size W x W x D and Dout number of kernels with a spatial size of F with stride S and amount of padding P, then the size of output volume can be determined by the following formula:

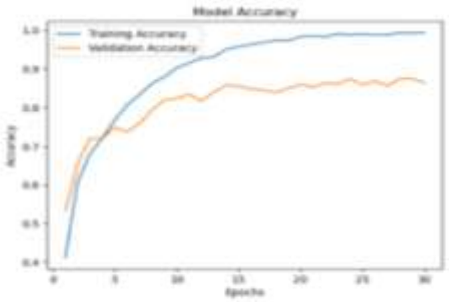
$$W_{out} = \frac{W - F + 2P}{S} + 1$$

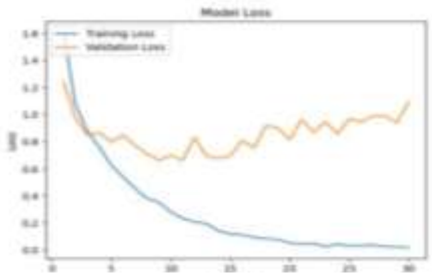
Pooling Layer:

The pooling layer replaces the output of the network at certain locations by deriving a summary statistic of the nearby outputs. This helps in reducing the spatial size of the representation, which decreases the required amount of computation and weights. The pooling operation is processed on every slice of the representation individually.

IX. MODEL ACCURACY

The function utilizes Matplotlib to create two separate plots: 1. The first plot displays the model's accuracy on the training and validation sets over the specified number of epochs. 2. The second plot shows the model's loss (typically a measure of how well the model is performing) on the training and validation sets over the specified number of epochs.





X. CONCLUSION

There are number of ways by which we can identify and classify the plants. Each has some pros as well as limitations. On one hand visual analysis is least expensive and simple method, it is not as efficient and reliable. The applications Neural Networks (NNs) have been formulated for classification of various plant leaves. Recognizing the features accurately and efficiently is mainly the purpose of the proposed approach. In this System, a deep learning convolutional neural network based on Keras and TensorFlow is deployed using python for image classification. The system works on the principle based on detecting a part of image and extracting CNN features from multiple convolutional layers. These features are aggregated and then given to the classifier for classification purpose. On basis of the results which has been produced, the system has provided the 95% accuracy in prediction of finding crop species.

XI. FUTURE SCOPE

In coming days, the mobile application can be developed to revolutionize plant identification by combining advanced machine learning techniques with multimedia integration. The core functionality involves not only accurately identifying plants based on visual cues but also automatically estimating the correctness of the identified species. By enhancing the model's robustness through an increased number of layers, the application seeks to ensure adaptability to diverse environmental conditions and a wide range of plant species. The user experience will be

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 02 Feb 2024, pp: 399-403 www.ijaem.net ISSN: 2395-5252

enriched with open multimedia content, including audio and video explanations, providing an engaging and informative platform for plant enthusiasts. Through meticulous data collection, annotation, and model training, the application aspires to set a new standard for reliable, user-friendly, and future-ready plant identification solutions.

REFERENCES

[1]. Dyrmann, Mads, H. Karstoft, and H. S. Midtiby. "Plant species classification using deep convolutional neural network." Biosystems Engineering

[2]. Jeon, Wang Su, and S. Y. Rhee. "Plant Leaf Recognition Using a Convolution Neural Network." International Journal of Fuzzy Logic & Intelligent Systems

[3]. Aparajita Sahay, Min Chen, “Leaf Analysis for Plant Recognition” Computing and Software System, School of STEM University of Washington Bothell

[4]. Akhil Arora, Ankit Gupta, Nitesh Bagmar, Shashwat Mishra, “A Plant Identification System using Shape and Morphological Features on Segmented Leaflets”

[5]. James S Cope, Paolo Remagnino, Sarah Barman, and Paul Wilkin “The Extraction of Venation from Leaf Images by Evolved Vein Classifiers and Ant Colony Algorithms” Digital Imaging Research Centre, Kingston University, London, UK

[6]. Yang, Li Wei, and X. F. Wang. "Leaf Image Recognition Using Fourier Transform Based on Ordered Sequence." International Conference on Intelligent Computing Springer Berlin Heidelberg

[7]. X. Wang, C. Du, X. Zhang, Y. Ma, G. Sun. “Plant Identification Using Deep Learning Models: A Comprehensive Review and Future Perspectives”

[8]. G. L. Barbedo. "Plant Image Analysis: Algorithms and Applications for a Smartphone-based System"

[9]. A. Delibaltov, H. Ogut. "Plant Species Recognition Using Convolutional Neural Networks"

[10]. D. A. Kumar, P. S. Chakravarthi and K. S. Babu, "Multiclass Support Vector Machine based Plant Leaf Diseases Identification from Color, Texture and Shape Features," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)

[11]. M. Y. Braginskii and D. V. Tarakanov, "Identification of plants condition using digital images," Journal of Physics: Conference Series

[12]. D. Bambil et al., "Plant species identification using color learning resources, shape, texture, through machine learning and artificial neural networks," Environment Systems and Decisions

[13]. R. A. Asmara, M. Mentari, N. S. Herawati Putri and A. Nur Handayani, "Identification of Toga Plants Based on Leaf Image Using the Invariant Moment and Edge Detection Features," 2020 4th International Conference on Vocational Education and Training (ICOVET)

[14]. B. Vijaya, "Leaf Recognition and Matching with Matlab," Thesis. Faculty of San Diego State University

[15]. S. Boran and I. Yücel, "Leaves recognition system using a neural network," 12th International Conference on Application of Fuzzy Systems and Soft Computing

[16]. K. Pankaja and V. Suma, "Mango leaves recognition using deep belief network with MFO and multi-feature fusion," Smart Innovation, Systems and Technologies