© JUN 2023 | IRE Journals | Volume 6 Issue 12 | ISSN: 2456-8880

Image Classification of Medicinal Flowers

BHAVYATHA B.[1], GAGANA R.[2], GHOWLI AMRUTHA[3], SAPNA R.[4]

[1] *Department of Information Science and Engineering, Presidency University, Bengaluru, India*

[2] *Department of Computer Science and Engineering, Presidency University, Bengaluru, India*

[3] *Department of Computer Engineering, Presidency University, Bengaluru, India*

[4] *Assistant Professor, Department of Computer Science and Engineering, Presidency University, Bengaluru, India*

Abstract- This research describes a method for classifying images of medicinal flowers using deep learning techniques. Since their therapeutic qualities have been used for generations, it is essential to correctly identify medicinal flowers before using them. The suggested technique starts with pre-processing the floral photos, then uses convolutional neural networks (CNNs) to extract features. A dataset of 10 distinct medicinal flowers was gathered and annotated with their respective classes in order to assess the performance of the suggested strategy. The study also looked into how various hyper-parameters, such the batch size, learning rate, and number of layers in the CNN model, affected classification accuracy. According to the findings, classification accuracy was increased by reducing the learning rate and adding more layers to the CNN model. The topic of plant identification and classification, particularly that of medicinal plants, has tremendous promise for applications of the suggested approach. This could help pharmacologists, researchers, and botanists identify and analyse various types of medicinal plants, ultimately resulting in the creation of more efficient cures.

I. INTRODUCTION

IMAGE: Light emitted from an item comes from a variety of angles, creating an image. How does a computer convert an image into a digital image, then? Pixels are small visual components that make up digital images. An array of rectangles is how these pixels are arranged. The number of columns in the array is used to compute the image's width, and the number of rows in the array is used to determine the image's height. Initially, electron beam scanning patterns on televisions were used to calculate digital images.

The scale of a digital image and the depiction of a real-world image can occasionally be misconstrued. The quantity of pixels determines the size of a digital image. One component called resolution is used to calculate the actual digital image. Resolution is a unique scale applied to image pixels. The units of resolution are pixels per inch (ppi), dots per inch (dpi), and lines per inch (lpi).

Now that the pixels have been established, our image only has a rectangular shape. Intensity is yet another factor that helps characterise an image. Each pixel has its own brightness, value, or intensity. The image would have a consistent tone if every pixel had the same value.

Bits make up data, which is divided into bytes and kilobytes. Kilobytes come after megabytes and are composed of 1024 kilobytes. We anticipate producing and processing a massive amount of data in the future, with a prediction of over 175 zettabytes in 2025. Global data generation is substantially influenced by images, such as photos and videos, and AI-based technologies like computer vision, deep learning, and machine learning may help organise and manage this data. Image classification, which employs Deep Learning models to analyse photographs and identify numerous features, from image content to the time of day, is one of the ways AI might be useful. This article will describe image classification, how it works, and how it may improve businesses.

II. CLASSIFICATION OF IMAGES

Image classification is the process of grouping and labelling sets of pictures or vectors in an image according to predetermined standards. A label might be given according to one or more criteria. There are two methods for classifying images: single-label and multi-label. Single-label assigns only one label to a picture, whereas multi-label assigns multiple labels to an image.

There are two different kinds of image classification techniques: single-label classification and multi-label classification. Each image is given a single label or category in single-label classification, based on a single criterion. An AI model may, for instance, be taught to classify photographs based on whether they depict daylight or night-time situations.

In contrast, an image can have many labels in multi-label categorization, and some photos may include all of the labels at once. Multi-label classification is frequently used to categorise movie posters, when a single film may fall under more than one category.

III. IMPORTANE AND SIGNIFICCAN OF MEDICINAL PLANTS

Introduction:

a description of medicinal plants and their use in conventional medical procedures
Overview of the use of medicinal herbs throughout history in various civilizations and cultures showcasing the various cultural meanings and practises with medicinal herbs

Medical Plants' Historical Relevance

Examining traditional medical systems from the past that extensively rely on medicinal plants, such as Ayurveda, Traditional Chinese Medicine, and indigenous healing methods

Examination of historical writings, archives, and manuscripts demonstrating the historical use of medicinal herbs

Case studies of key medicinal plants that were instrumental in the development of historical medical procedures

IV. WORKINGS OF IMAGE CLASSIFICATION

Pre-processing, object detection, object recognition and training, object classification, and linking to an AI workflow are all steps in the process of classifying images. Pre-processing include clearing the data and making sure that all the photos are high-quality and pertinent. Finding items within an image set is the task of object detection, whereas labelling found things, seeing patterns, and training an AI model are the tasks of object recognition and training. The AI model categorises the photos in the last step of object categorization using several criteria. The AI workflow that specifies the input and output of the data can be connected to the AI model at this point.

V. IMAGE CLASSIFICATION TECHNIQUES

The preparation of photos of medicinal plants before classification requires the use of image processing tools. Image quality is improved by using noise reduction techniques like median filtering or Gaussian filtering to remove undesired pixel changes. Techniques for improving images include histogram equalisation and contrast stretching, which adjust brightness and increase visual details. Segmentation techniques isolate the areas with medicinal plants from the surrounding areas, allowing for easier analysis. Effective categorization is made possible by feature extraction algorithms, which extract pertinent features from the photos, such as texture, shape, or colour information. These image processing methods work together to improve the quality, extract key details, and segment the regions of interest, which helps to accurately categorise photos of medicinal plants.

- Machine learning algorithms

For classification tasks, such as the classification of medicinal plants, machine learning methods are frequently used. Support vector machines (SVM) can handle high-dimensional data and provide efficient class separation utilising hyperplanes. An ensemble of decision trees is used by random forests to provide robustness against overfitting and the capacity to handle big datasets. K-nearest neighbours (KNN) classify samples based on how close they are to instances that have been given labels, and they are very helpful for non-linear data. Decision trees can handle categorical and numerical data and offer interpretability. Each technique, however, has advantages and disadvantages, such as the SVM's sensitivity to parameter adjustment or the computational expense of KNN for large datasets. Selecting the best algorithm for classifying medicinal plants is aided by an understanding of these subtleties.

- Deep learning methods:

Due to its capacity to automatically learn complicated properties from photos, deep learning techniques, notably convolutional neural networks (CNNs), have drawn a lot of interest in the classification of medicinal plants. CNNs are highly suited for analysing the complicated features of photos of medicinal plants because they are excellent at capturing hierarchical patterns and spatial relationships within images. CNNs may extract discriminative features and achieve high classification accuracy by utilising numerous convolutional layers, pooling layers, and fully connected layers. The classification of medicinal plants has undergone a revolution thanks to the ability to learn from unprocessed visual data without intentional feature engineering, allowing for more precise and effective identification and characterization of plant species.

Utilising insights gained from other picture classification tasks, transfer learning and pre-trained models have become effective methods for classifying medicinal plants. Transfer learning allows for the transfer and fine-tuning of learned representations from a source task to a target goal, such as the classification of medicinal plants. Pre-trained models give a head start by capturing general visual traits and are pre-trained on massive image datasets like ImageNet.

There are various benefits to using transfer learning and pre-trained models when classifying medicinal plants. The considerable labelled data from the source task is used to first get around the restriction of little labelled data in the medicinal plant domain. As a result, even with a smaller labelled dataset of medicinal plants, the model can acquire strong and generalizable characteristics. Second, because the early layers have already acquired basic visual properties, transfer learning and pre-trained models dramatically minimise the training time and computational resources needed.

The model can now capture domain-specific properties and nuanced characteristics of medicinal plants thanks to transfer learning and pre-trained models, which simplify the transfer of knowledge from related tasks. As a result, classification accuracy and generalisation performance are enhanced. Additionally, they offer a framework for continuous learning, allowing the model to be improved upon or modified using fresh information about medicinal plants. However, it is crucial to take into account the domain variations between the source and target tasks, as some visual characteristics might not be immediately applicable to the classification of medicinal plants. To overcome these difficulties and improve the transfer learning procedure, one can use fine-tuning approaches, selective layer freezing, or domain adaption techniques.

Generally speaking, transfer learning and pre-trained models have shown their efficacy in enhancing the effectiveness, accuracy, and generalisation capabilities of medicinal plant classification models, offering a useful strategy in leveraging existing knowledge and accelerating the development of robust classification systems in this field.

- What key concepts underlie image classification?

Let's look at some key words and technologies to better grasp the training procedure and how picture classification functions.

Here is a quick summary before we dig into each of these ideas in more detail:

The data and its structure decide the type of machine learning (supervised or unsupervised).

The process can only be optimised by training on high-quality datasets.

Similar to human vision, AI-powered computer vision enables machines to recognise things in images.

Let's now investigate the main factors that affect image classification.

Unsupervised learning and supervised learning are the two methods used in machine learning. Large volumes of data must be collected in order to teach a machine how to identify photos.

The more common strategy, supervised learning, involves people giving computers labelled example data with the right answers. This makes it possible for the machine to identify correlations and use them with fresh data. Unsupervised learning, on the other hand, is characterised by raw, unstructured data without human interaction. This approach uses deep learning but does not utilise training data. Unsupervised learning, on the other hand, can produce insights that humans have not yet recognised.

- Algorithms for image classification

Machine learning or deep learning are two categories in which image categorization algorithms can be categorised. Despite similarities, there are some distinctions that affect image classification.

Machine learning makes use of algorithms that, by learning from data, can complete tasks without human intervention.

The processing of unstructured data, such as text, photos, and documents, uses algorithms inspired by the workings of the human brain in a technique called deep learning. Similar to how neurons in the brain communicate information between nodes, it depends on neural networks. Each node analyses data and transmits its findings to the nodes in the layer below it. In response, a non-linear abstraction of the data is made. There are several types of Neural Networks, including Convolutional Neural Networks (CNNs), which are used extensively for image recognition and processing. In CNNs, the output of nodes in the hidden layers isn't always shared with every node in the next layer, making it ideal for identifying objects in photos.

- What is a convolutional neural network?

A convolutional neural network (CNN) is a type of artificial neural network used primarily for image recognition and processing, due to its ability to recognize patterns in images.

Due to its capacity to spot patterns in images, a convolutional neural network (CNN) is a type of artificial neural network that is mostly used for image recognition and processing.

Let's break things down into clear terms.
Your brain produces numerous assumptions when you see a picture of a house, as if it were a drawing or picture of a house that just naturally and effortlessly occurred to you. We come to the conclusion that it depicts a dwelling. Now, a computer does not naturally possess the ability to conduct object identification that we do. Here is where the CNN (Convolutional Neural Network) idea can be put to use. Deep learning has a section dedicated to pattern recognition.

VI. EVALUATION METRICS AND PERFORMANCE ANALYSIS

When evaluating the effectiveness of categorization models for medicinal plants, evaluation measures are essential. Insights into the model's accuracy, precision, recall, F1-score, and overall performance are provided via a number of widely used metrics.

When classes are balanced, accuracy—which counts the percentage of occurrences that are correctly classified—is appropriate. Indicating the model's capacity to prevent false positives, precision is the percentage of accurate positive predictions among all positive predictions. Recall, often referred to as sensitivity, gauges how many true positive predictions are made out of all real positive examples, demonstrating the model's capacity to prevent false negatives. The harmonic mean of recall and precision, known as the F1-score, provides a balanced indicator of model performance. Confusion matrix analysis provides a thorough evaluation of classification performance in addition to these indicators. It enables in-depth investigation of particular model faults by providing the number of true positives, true negatives, false positives, and false negatives. Metrics like precision, recall, and accuracy can be generated from the confusion matrix.

In addition, different specialised evaluation metrics may be used in accordance with the demands of the work of classifying medicinal plants. These can include average precision (AP) for imbalanced datasets or area under the receiver operating characteristic curve (AUC-ROC).

Knowing and examining these evaluation criteria can shed light on the benefits and drawbacks of categorization algorithms for medicinal plants, directing future development and ensuring peak performance.

- How cnn works

Let's now look at an example of a typical network that has several interconnected layers. Each layer now gets an input, which is transformed into something else, and passes output to the following layer. Consider a part of the layers; within these layers, there are things referred to as filters that aid CNN in pattern detection. Take the home as an example; if it were a genuine image, it would be made up of a number of pixels. When we enlarge a specific area of the image, let's say a window, we may argue that the window is constructed entirely of perfectly straight lines. However, not every window is constructed with completely straight lines. Some also feature curved lines. The cool property of filters that makes CNN so excellent is this.

We can define a pattern to check for within the 3*3 block that makes up the filter.

VII. THE PROPOSED SYSTEM'S MODEL

First, we spent the majority of the time collecting data by taking pictures of uncommon medicinal plants' blossoms and leaves. After the data was gathered, a model was developed using CNN to train each of the leaf and flower images. The input picture, the feature detector or mask, and the output image are the three parts that make up a CNN convolution. The mask is taken in 3x3 format for precision. The 3x3 mask and input picture are used to extract the image's features. The input image is covered by the 3x3 mask, as can be seen in the figure below. The highest value pixel will be extracted by constantly applying the 3x3 mask to the input image by overlapping it. A 2x2 image will be produced, as seen in Figure. Some aspects of the supplied image will be ignored throughout this process. However, the plant's identification can be done more accurately by isolating the pixels with the highest value. The trained photographs will be kept in the database so that the identification process can continue. The trained photos will be put to the test using the predictions during the testing phase, with the goal of recognising the uncommon medicinal plant. The number of photos and the number of epochs will have an impact on the accuracy. Epoch is a count of how many times the training has been carried out. High accuracy can be attained after the epoch count is between 40 and 60.

VIII. CHALLENGES

The classification of medicinal plants faces a number of difficulties and constraints that affect the precision and potency of the models. Several of these difficulties include:

Data imbalance: Medicinal plant datasets frequently have unbalanced class distributions, meaning that some plant species have a disproportionately high number of samples compared to others. As a result, models may be biassed and perform well for dominant classes while struggling for minority classes.

Inter-class Variability: Because of things like growth circumstances, environmental influences, and genetic differences, medicinal plants differ significantly both within and between species. The considerable inter-class diversity makes it difficult to detect and precisely describe the distinctive characteristics of many plant species.

Scalability: As the number of medicinal plant species rises, it becomes more difficult to scale categorization models. managing and instructing models

IX. FUTURE DIRECTIONS

Multimodal Classification: Integrating several modalities, such as merging textual and spectral data with picture data, can offer further insights for the classification of medicinal plants. This combination of data sources can increase the accuracy and discriminative capability of models.

Integration of Domain Knowledge: The feature extraction and classification process can be guided by the integration of domain knowledge, such as botanical knowledge, conventional medical procedures, or phytochemical qualities. The performance of classification can be improved and model interpretability can be improved by incorporating prior information.

Data Synthesis, Data Augmentation, and Transfer Learning: Increasing the training data through these methods can help lessen the effects of data imbalance and inter-class variability, allowing models to generalise more effectively.

X. APPLICATIONS

The correct classification of medicinal plants has a variety of useful uses and advantages that have a big impact on many different disciplines. Several of the main uses and advantages include:

Drug Discovery: The discovery of plant species with potential therapeutic characteristics is aided by the precise taxonomy of medicinal plants. The discovery and development of novel pharmaceuticals and organic compounds for a range of illnesses and ailments can be aided by this understanding.

Formulation of Herbal Medicine: Correct classification enables the formulation of herbal medicines with certain plant combinations and proportions, ensuring the efficacy and security of conventional treatments. It allows for the standardisation of herbal compositions and promotes the use of traditional medicine's evidence-based procedures.

Correct classification aids in the preservation and conservation of rare or endangered medicinal plant species. Conservation efforts can be directed towards maintaining and sustaining these plants' populations, ensuring their availability for future generations, by finding and comprehending their medical benefits.

Practises of Traditional Medicine: The classification of medicinal plants helps to record and preserve traditional practises and knowledge. By presenting scientific proof and comprehending the active ingredients and processes of action in conventional treatments, it contributes to the validation of traditional medical systems.

Sustainable Agriculture: Correct classification makes it easier to cultivate and manage medicinal plants sustainably. It assists in deciding which plant species are most suited for cultivation, enhancing growing conditions, and verifying the purity and effectiveness of plant-based therapeutic products.

CONCLUSION

In conclusion, Neural Networks (NN) have demonstrated encouraging results in the picture categorization of medicinal flowers. We can train NN models to correctly classify photos of diverse medicinal flowers by utilising the power of deep learning, which can be helpful for pharmaceutical businesses, researchers, and herbal medicine practitioners. Convolutional Neural Networks (CNNs) in particular have demonstrated excellent performance in object identification and image processing methods, making them a perfect fit for this kind of application. We may anticipate many more developments in the field of picture categorization using NN as technology develops, which will lead to improvements in the study and use of medicinal flowers.

REFERENCES

[1] Govaerts, R., Frodin, D.G., & Bogner, J. (2001). *World Checklist and Bibliography of Araceae (and Acoraceae)*. Royal Botanic Gardens, Kew.

[2] Duke, J.A., & Ayensu, E.S. (1985). *Medicinal Plants of China. Reference Publications*.

[3] Farnsworth, N.R. (1988). *Screening Plants for New Medicines. In Wilson, E.O. (Ed.), Biodiversity*. National Academy Press.

[4] Heinrich, M., & Gibbons, S. (2001). *Ethnopharmacology in Drug Discovery: An Analysis of its Role and Potential Contribution. Journal of Pharmacy and Pharmacology*, 53(4), 425-432.

[5] Fabricant, D.S., & Farnsworth, N.R. (2001). *The Value of Plants Used in Traditional Medicine for Drug Discovery. Environmental Health Perspectives,* 109(Suppl 1), 69-75.

[6] WHO (World Health Organization). (2002). *Traditional Medicine: Report by the Secretariat. Fifty-sixth World Health Assembly*.

[7] Bensky, D., & Gamble, A. (1993). *Chinese Herbal Medicine: Materia Medica. Eastland Press*.

[8] Arora, R. (2013). Medicinal Plants Classification: *An Introduction. In Medicinal Plants Classification, Identification, and Utilization. CRC Press*.

[9] De Luca, V., Salim, V., & Thamm, A.M.K. (2012). *Mining the Biodiversity of Plants: A Revolution in the Making.* Science, 336(6089), 1658-1661.

[10] Sharma, V., Sharma, A., & Kansal, L. (2019). *Medicinal Plants: Source of New Drugs. International Journal of Pharmaceutical Sciences and Research*, 10(1), 1-11.