Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Medicinal Plant Recognition Based on the Seedling Image and Deep Learning

Khalid M. O. Nahar[1], Nesrine Atitallah[1], Mohammed M. Abu Shquier[2,*], Dalal Z. Zreiqat[3], Khaled M. Alhawiti[4]

[1] Faculty of Computer Studies, Arab Open University, Hittin, Riyadh 11681, Saudi Arabia
[2] Computer Science Department, Faculty of Computer Science and Information Technology, Jerash University, Jerash, Jordan
[3] Computer Science Department, Faculty of IT and Computer Sciences, Yarmouk University, Irbid 21163, Jordan
[4] Computer Science Department, Faculty of Computers and IT, University of Tabuk, Tabuk 47512, Saudi Arabia

ABSTRACT

In the realm of botanical science and environmental conservation, identifying medicinal plant species accurately and efficiently is crucial for their sustainable cultivation and conservation. This paper introduces a novel approach that leverages seedling images and deep learning techniques to develop a computer vision system. The system aims to recognize medicinal plants at their early growth stages, a task that is fundamental for effective plant management and preservation efforts. The proposed work utilizes the Inception 3, a deep convolutional neural network (CNN) architecture, trained on a comprehensive dataset of seedling images from diverse medicinal plant species. Evaluating the system's performance with various metrics reveals its exceptional accuracy, reaching up to 97.6% in identifying medicinal plant species. This model holds promise for applications in plant conservation, biodiversity assessment, and the cultivation of medicinal plants.

Keywords:
Deep learning; convolutional neural network; medical plants; seedling images

1. Introduction

Throughout history, humans have sought ways to overcome the many diseases that threaten their health and well-being. From the dawn of his existence, man has continuously battled diseases, turning to herbs and medications as his primary means of protection. Despite the antiquity of recorded data on therapeutic plants, archaeological research has revealed that the Sumerians documented the proven therapeutic use of plants such as laurel, caraway, and thyme at least 5,000 years ago. Furthermore, evidence suggests that the use of herbal remedies dates back to at least 60,000 years ago in Iraq and 8,000 years ago in China [1].

Recently, the use of herbs has increased due to the side effects of chemical drugs, lack of effective treatments for chronic illnesses, and microbial resistance despite record-breaking pharmaceutical research expenditures. Aromatic medicinal herbs, known for their strong aroma and unique flavour,

* *Corresponding author.*
E-mail address: shquier@jpu.edu.jo

https://doi.org/10.37934/araset.58.1.209219

are used both for seasoning food and their medicinal properties. These herbs improve food quality by enhancing taste, preservation, and providing anti-oxidation and anti-mildew benefits due to their volatile compounds.

The utilization of therapeutic plants has significantly grown alongside advancements in medical science research. Jordan's diverse topography and climate have resulted in a wide variety of vegetation, particularly wild plant species. Identifying plant seedlings is crucial for distinguishing between species. Comprehensive knowledge of plant structures, such as petals and leaves, is essential for monitoring environmental conditions and modelling climate change using machine learning techniques [2,16].

Additionally, there is significant interest in image classification. It is critical to develop a quick system for classifying plant species given the accessibility of digital photos from online sources. CNN expands its reach and develops the capacity to implement original solutions to a variety of problems, including the classification of plants in agricultural fields. Many scholars have been working on developing automated analysis of plant photos. Recently, a review of the various methods employed was also presented. They proved to be practical for increasingly complex, multifaceted picture identification problems as a result. Currently, CNNs are used in a variety of modernistic picture-sorting jobs that can handle extremely complex image identification with several entity classes to an astounding level. We propose in this research deep learning-based convolutional neural networks for classifying plant seedlings in photos. On the basis of accuracy, specificity, and sensitivity tests performed on photos of plant seedlings, the plant classification model is assessed. The findings of this study will significantly advance crop management methods and contribute to the development of a new tool to facilitate and simplify agricultural procedures.

The paper is structured as follows: Section 2 explores the concept of deep learning, plant classifications, applications, and challenges. Section 3 details the proposed method for deep learning-based plant image classification. Section 4 provides a summary of the experiments and their results. Finally, Section 5 presents the conclusions.

2. Literature Review

Plant classification has been a pivotal area of study, leveraging various machine learning techniques to enhance the accuracy and efficiency of identifying and categorizing plant species, as depicted in Figure 1. One of the most commonly used strategies for classifying plant leaves involves employing machine learning and deep learning models. Neural networks, such as Artificial Neural Networks (ANN), Probabilistic Neural Networks (PNN), and Convolutional Neural Networks (CNN), are frequently utilized.

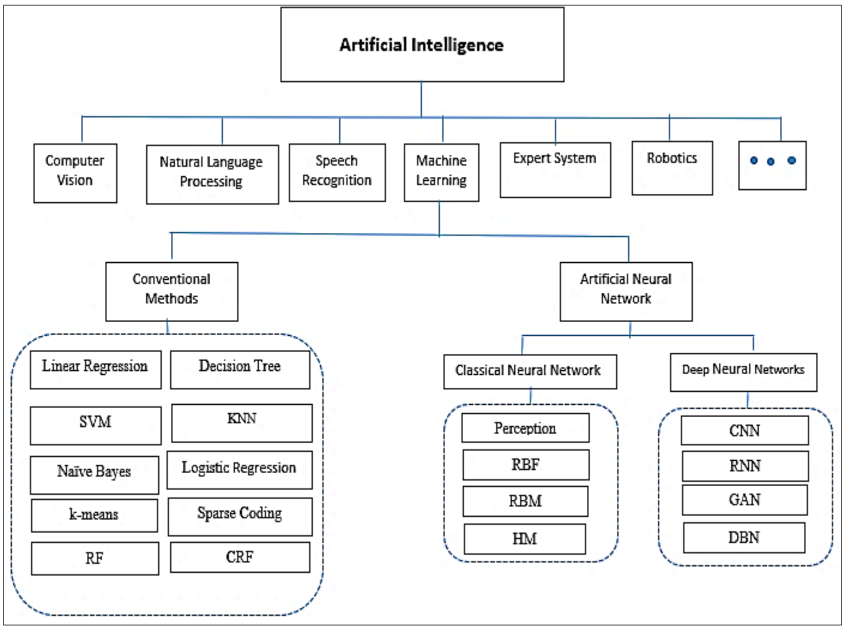

Fig. 1. Sample augmentation angles

Some studies have even integrated multiple techniques to increase accuracy. It revealed that the efficiency of plant leaf classification might be enhanced by employing a number of distinct pre-processing techniques and distinctive factors in feature extraction [9]. The deep neural network (DNN), also known as deep learning DL is a method that aims to incorporate "depth" (complex) behaviours and can address issues that need intricate, highly variable functions [4,8]. Based on DL neural network training algorithms typically use very large, and frequently unlabelled, data sets. A lot of studies use a dataset in various developmental stages to demonstrate a method for classifying plant seedlings. For image recognition, Convolutional Neural Network (CNN) algorithms a deep learning technique were used [3,5,6]. The authors take advantage of the strength of deep CNNs for automatic joint features and classifier learning. Additionally, they take advantage of the LSTMs' potential to research plant growth and their dynamic behaviours as crucial discriminative characteristics for accession categorization. The authors in other studies concentrated on the special task of classifying urban flora. The work's potential practical use is a tool that helps people who are growing plants at home identify new species and provide the necessary care instructions. The study presents Urban Planter, a novel dataset for classifying plant species that contains 1500 photos divided into 15 categories [1]. In another study [2], a smartphone application is developed that can identify a plant by taking a picture of it and then processing the image to determine when the soil should be changed, how much sunlight it needs, and what nutrients the plant needs. Convolutional neural networks are used to train the model, and the dataset is successfully included in the network. The authors in the other papers suggested a useful and automatic classification system to identify Malaysian herbs that would be employed in cooking or medicine. Which was then connected with a mobile app to make real-time classification easier. The suggested system used the Support Vector Machine (SVM) and Deep Learning Neural Network as its two classifiers (DLNN). The experimental results demonstrated that for both the experimental model and the created mobile app, SVM obtained 74.63% identification accuracy and DLNN reached 93% recognition accuracy. Additionally, the processing time using the mobile app was only 2 seconds, compared to 4 seconds for SVM and 5 seconds for DLNN classifier [4]. Within some others studies they used two methods—the standard method and the deep learning approach—are used to identify plant species. Conventional feature extraction makes use of Humoments (shape features), Haralick textures, local binary patterns (LBP), and colour channel statistics (colour features). The retrieved characteristics are classified using several classifiers, such as linear discriminant analysis, logistic regression, a classification and regression tree, naive Bayes, a k-nearest neighbour classifier, a random forest classifier, and a bagging classifier.

Additionally, various deep learning architectures are evaluated in the context of identifying plant species [7]. While other authors focused on the issue of insufficient training data or an uneven class balance in the datasets, which is the issue that is most brought up in the field of machine learning. This issue can be solved in part by a process known as data augmentation. They have compared and examined various data augmentation techniques for the task of classifying images in the paper, Style Transfer and Generative Adversarial Networks, together with the representative, are at the further end of a spectrum that begins with the more conventional approaches of image manipulation, such as rotation, cropping, zooming, and histogram-based methods [10].

Traditional medicine is widely used in some countries like China, India, Japan, Pakistan, Sri Lanka, and Thailand. China estimates that traditional tribal remedies account for about 40% of all medical consumption. Lentils found in the Caesalpiniaceae, Fabaceae, and Mimosaceae are used in Thai herbal treatments. The sales of herbal remedies are thought to have generated more than US$2.5 billion in the middle of the 1990s. In Japan, herbal pharmaceutical items are more in demand than conventional pharmaceuticals. Plants play a significant role in a variety of industries, including fine chemicals, cosmetics, pharmaceuticals, and industrial raw materials, among others. A vital role played by medicinal plants in the discovery of new drugs. The only effective means of treating certain conditions has been demonstrated to be medicinal plants. It is also India that has frequently been referred to as the Medicinal Garden of the globe since nature has endowed our nation with an enormous wealth of medicinal plants [20].

The medical application of plants mentioned in the Indian Vedas to treat various ailments. The traditional medical system is currently extensively accepted and used by people all around the world. Ayurveda, Siddha, Unani, Homeopathy, Yoga, and Naturopathy are some of the recognized Traditional systems of medicine that are currently practiced in India. Due to their drug-like qualities, medicinal plants have been recognized as possible medication candidates [12]. According to WHO, herbal remedies are crucial parts of primary healthcare, and approximately 11% of people worldwide use them of the 252 medications come from plants. High concentrations of numerous "super nutrients," including disease-fighting phytochemicals, antioxidants, and micronutrients, can be found in vegetarian diets.

Plants can act as antioxidants, antiviral, anticancer, antibacterial, antifungal, and antiparasitic, among other pharmacological functions. Plants include compounds that behave like antioxidants and scavenge free radicals, such as flavonoids, phenolics, anthocyanins, and vitamins. It has been suggested that phytochemicals' antioxidant properties may help to reduce oxidative stress in the biological system. Numerous human diseases, such as cancer, diabetes, hepato-renal disease, cardiovascular disease, and neurological disorders have seen a decrease in risk because of phytochemicals. But many herbal medications are made either directly or indirectly from plants that are regarded as essential modern medicines for treating a range of human ailments [13,18].

Table 1 summarizes the literature clarifying the dataset used, the approach, the features, the classifier used, and the achieved accuracy.

Table 1
Summary of the literature

Ref.	Dataset	Approach	Features	Classifiers	Accuracy
[2]	1500 images categorized into15 categories	plant classification	They used manually defined features	supervised end-to-end multiple DNNs	94–96% classification accuracy reached only 6.67%
[3]	several images of 43 different plants taken from various directions	- Convolutional Neural Networks - classified plant	leaf vein patterns	Deep learning techniques	86.2% success
[4]	public dataset of 4234 plant images	- hand-crafted approach - classification of herbal leaves	the shape of leaves, colour, and texture through ant colony optimization	- Template Matching - ANN - SVM - DBN	99.74% during validation 99.69% during testing.
[5]	leaves	-Classification Approach Based on Shape and Texture Features -End-to-End DL	- GLCM - ASM - IDM - entropy -correlation	- SVM - DLNN	SVM RBF kernel = 74.63% DLNN = 93.00%
[6]	5608 images with 960 unique plants	Nonlinear approach	added data like a retailer, sender, user, social media event, credit score, IP address, and a host of other features	-Neural Network (CNN) algorithms. -a deep learning technique extensively applied to image recognition was used	99.48% on a held-out test set
[7]	It is a subset of a much larger ImageNet dataset and has about 1.2 million images selected from 1000 different categories	temporal matching approach	traditional hand-crafted image analysis features and other accession classification frameworks	CNN-LSTM framework	96%
[8]	BJFU100 dataset is used and it consists of 10000images of 100 ornamental plant species found in Beijing Forestry University campus	Deconvolution Network (DN) approach	-leaf features are extracted based on shape and texture. -Hybrid Feature Selection (HFS)	-LDA -LR -NB -KNN -CART -RF -BC	99.65%
[9]	2171 pictures taken from ImageNet	picture classification		- DNN	90%
[10]	Flavia dataset	high performance statistical approaches have been used to perform leaf features	Leaf features such as shape, size, and colour. The contour-based extraction is expressed in length, width, aspect ratio, and leaf diameters descriptors	- ANN - CNN - PNN - SVM	74.5% using the fusion features. 59.6% and 50.8%, respectively, without fusion
[11]	more than one million images from 1000 non-overlapping categories	A classical approach	based on a combination of affine image transformations and colour modification	deep learning GANs	N/A

3. Proposed Methodology

This section outlines the approach adopted to develop a reliable and efficient solution for the early-stage identification of medicinal plant species, which has significant implications for plant conservation and medicinal plant cultivation. The proposed methodology, illustrated in Figure 2, provides a comprehensive workflow, with data augmentation focused on generating multiple views of the image from various angles. The upcoming sections detail the proposed methodology [17].

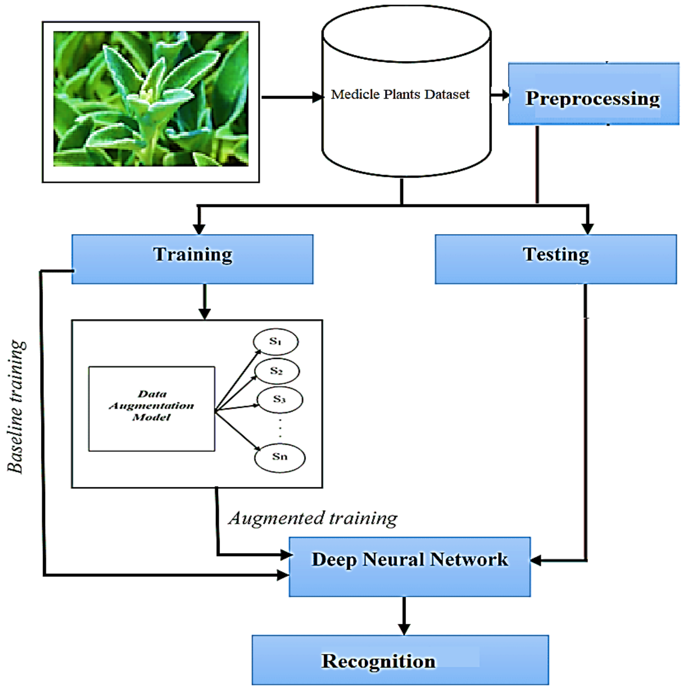

Fig. 2. Methodology flow diagram

3.1 Data Collection

Most of the dataset for this publication comes from numerous photographs. These pictures Some of them were taken from photographs in nature, while others were taken from Google Images. There were numerous images of plants in this study. This research paper's exclusive focus is on how to incorporate medicinal plants into all its classes. There are a lot of plant photographs, and there are eight different kinds of plants included. Numerous images of various plant species are available, each with a unique aspect and colour. The total number of images for this bloom is 4030, as shown in Table 2. Each image is resized to a 250x250 pixel size, and the RGB (red, green, and blue) values of the pixels are recorded in the dataset. Input shape is created as a result (250, 250, 3). 20% of the dataset is used for testing and validation, and the remaining 60% is used for training. For each plant species, a training dataset construction was picked at random. Most of the visuals were produced using stock photos. Shots taken manually were made. Many of these handheld images were taken in well-lit settings, while others were taken in inside lighting.

Table 2
Types of plants used in dataset

Sl.no	Type of Plants	Arabic Name	No of Images	Sample
1	Caraway	كراوية	517	
2	Chamomile	بابونج	500	
3	Laurel	ورق غار	540	
4	Mint	نعناع	523	
5	Rosemary	اكليل الجبل	505	
6	Rue	فيجل	500	
7	Sage	ميرمية	518	
8	Thyme	زعتر	427	

3.2 Data Preprocessing

The collected seedling images are pre-processed to remove any noise or artefacts and to enhance the features that are important for plant recognition. This step may involve image resizing, cropping, filtering, and normalization [19].

3.3 Data Augmentation

The pre-processed seedling images are augmented to increase the diversity of the dataset and improve the generalization ability of the model. This step may involve random cropping, flipping, rotation, and adding noise. Figure 3 shows the sample augmentation angles.

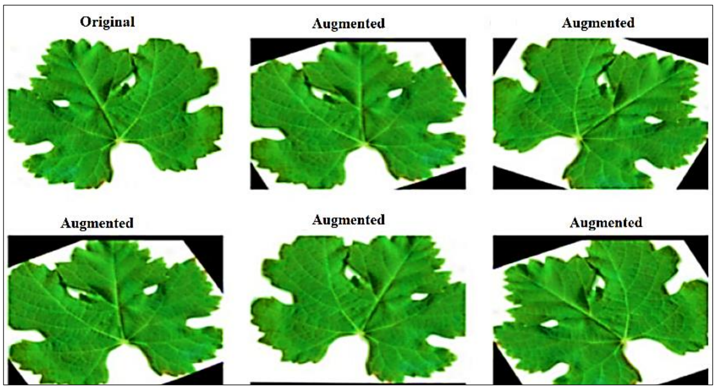

Fig. 3. Sample augmentation angles

3.4 Training Model

A deep convolutional neural network (CNN) architecture is designed and trained on the augmented dataset. CNN should have multiple layers of convolution, pooling, and activation functions, followed by a fully connected layer and a SoftMax output layer. The training process involves minimizing the cross-entropy loss between the predicted and actual plant species labels. CNN networks and deep learning are utilized to create the model that will be used for plant identification and information provision in a mobile context. CNN networks are specialized networks that are utilized in the processing of picture data. LeCun *et al.,* first introduced CNN in 1990 [15]. In order to classify images, determining factors are used. These could include, for instance, plant leaf side designs. The filters of a CNN structure are one of its key advantages. As a result, image classification may be done effectively regardless of where it is. This would save on an absurd number of weights. CNN networks are made up of several layers. The convolution layer is one of them and is crucial. In CNN, four crucial criteria are applied. These are the following: filter size, and no padding.

The third iteration of a series of Deep Learning Convolutional Architectures, Inception V3, was the one we used. Inception V3 was trained with a dataset of 1,000 classes derived from the original ImageNet dataset, which was trained using more than a million training images. The Tensor Flow version of Inception has 1,001 classes due to an additional "background" class not being used in the original ImageNet. Trained for the ImageNet Big Visual Recognition Competition, Inception V3 finished in second place overall. Figure 4 is a pictorial representation of the inception 3 model for medical plant recognition.

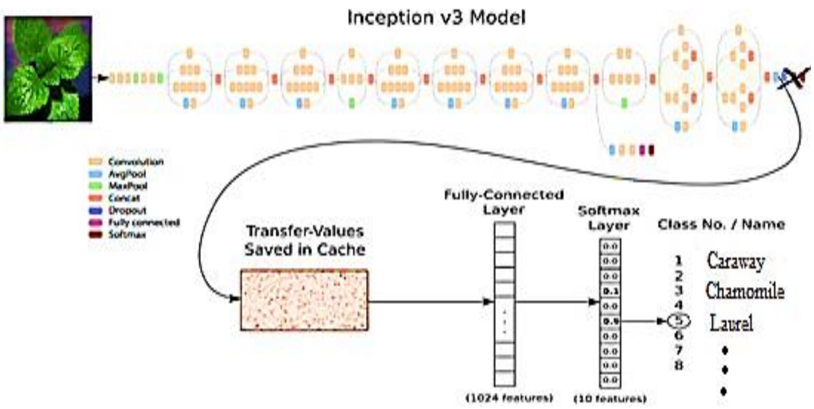

Fig. 4. Inception V3 model

3.5 Model Evaluation

The trained CNN is evaluated on a test dataset to assess its performance in terms of accuracy, precision, recall, and F1 score. The evaluation metrics are used to compare the proposed model with other state-of-the-art methods. We focus on accuracy measurement.

3.6 Model Deployment

The trained CNN model is deployed in a real-world scenario to recognize medicinal plant species from seedling images. This may involve developing a user-friendly interface or integrating the model with a mobile application.

4. Results

The proposed model was trained on 8 types of medical plants, the model ran for 10, 15, 20, 30, 40, 50, 60, 80, and 100 epochs. The accuracy for each species is evaluated and the average accuracy was also declared. Table 3 shows the accuracy and the validation loss for each plant type. The overall accuracy was 97.4%, which is considered promising.

Table 3
Results for different scenarios

Class name	Accuracy	Val Loss
Caraway	0.9633	1.03385
Chamomile	0.9811	1.6011
Laurel	0.9753	1.4782
Mint	0.9610	1.01728
Rosemary	0.9629	1.0030
Rue	0.9773	1.5401
Sage	0.9795	1.5976
Thyme	0.9923	1.7602
Average	0.9740875	1.37891625

A pictorial view of Table 3 is presented in Figure 5; it is clear that the accuracy is almost the same for all species with some little variability due to the number of training samples.

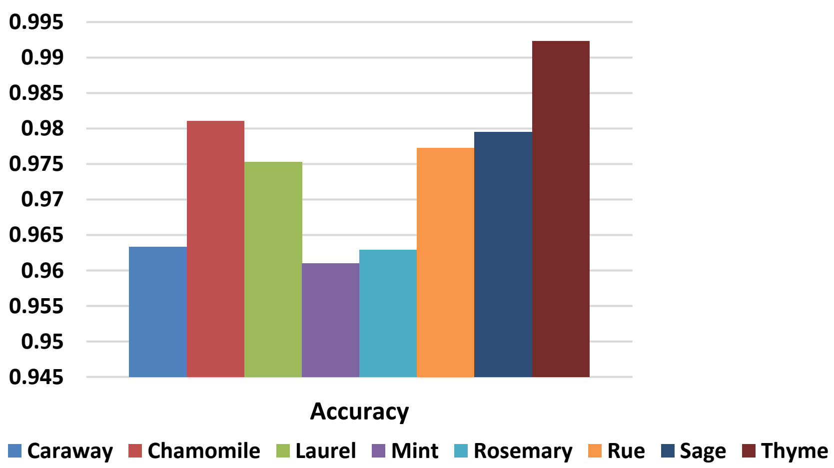

Fig. 5. Accuracy evaluation for each species

The validation loss of each plant type is also presented in Figure 6.

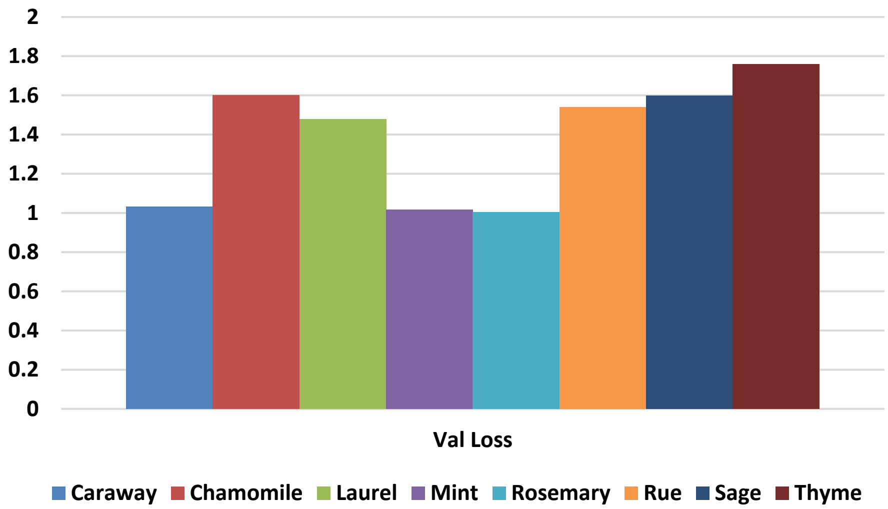

Fig. 6. Validation Loss for each species

5. Conclusions and Future Works

In every industry, the proliferation of smart systems is making life easier, with deep learning techniques playing a crucial role in creating these sophisticated systems. This work focuses on developing a model to identify plant types from images, offering the potential to simplify understanding plant requirements in social contexts. The tests and model achieved high levels of accuracy, demonstrating the effective use of Convolutional Neural Networks (CNN) in processing images of this nature and similar ones.

Enhancing the model's accuracy rate can be achieved by training the network with new patterns, a process simplified with CNN. Future research can further improve the model by expanding the dataset and adding more classes.

Acknowledgments

The authors extend their appreciation to the Arab Open University for funding this work through AOU research fund No. (AOUKSA-524008).

References

[1] Pan, Si-Yuan, Gerhard Litscher, Si-Hua Gao, Shu-Feng Zhou, Zhi-Ling Yu, Hou-Qi Chen, Shuo-Feng Zhang, Min-Ke Tang, Jian-Ning Sun, and Kam-Ming Ko. "Historical perspective of traditional indigenous medical practices: the current renaissance and conservation of herbal resources." *Evidence-Based Complementary and Alternative Medicine* 2014, no. 1 (2014): 525340. https://doi.org/10.1155/2014/525340

[2] Litvak, Marina, Sarit Divekar, and Irina Rabaev. "Urban plants classification using deep-learning methodology: a case study on a new dataset." *Signals* 3, no. 3 (2022): 524-534. https://doi.org/10.3390/signals3030031

[3] Adak, Muhammed Fatih. "Identification of plant species by deep learning and providing as a mobile application." *Sakarya University Journal of Computer and Information Sciences* 3, no. 3 (2020): 231-238. https://doi.org/10.35377/saucis.03.03.773465

[4] Alimboyong, Catherine R., Alexander A. Hernandez, and Ruji P. Medina. "Classification of plant seedling images using deep learning." In *TENCON 2018-2018 IEEE Region 10 Conference*, pp. 1839-1844. IEEE, 2018. https://doi.org/10.1109/TENCON.2018.8650178

[5] Muneer, Amgad, and Suliman Mohamed Fati. "Efficient and automated herbs classification approach based on shape and texture features using deep learning." *IEEE Access* 8 (2020): 196747-196764. https://doi.org/10.1109/ACCESS.2020.3034033

[6] Ashqar, Belal AM, Bassem S. Abu-Nasser, and Samy S. Abu-Naser. "Plant seedlings classification using deep learning." (2019).

[7] Taghavi Namin, Sarah, Mohammad Esmaeilzadeh, Mohammad Najafi, Tim B. Brown, and Justin O. Borevitz. "Deep phenotyping: deep learning for temporal phenotype/genotype classification." *Plant methods* 14 (2018): 1-14. https://doi.org/10.1186/s13007-018-0333-4

[8] Anubha Pearline, S., V. Sathiesh Kumar, and S. Harini. "A study on plant recognition using conventional image processing and deep learning approaches." *Journal of Intelligent & Fuzzy Systems* 36, no. 3 (2019): 1997-2004. https://doi.org/10.3233/JIFS-169911

[9] Kothari, Jubin Dipakkumar. "A case study of image classification based on deep learning using TensorFlow." *Jubin Dipakkumar Kothari (2018). A Case Study of Image Classification Based on Deep Learning Using Tensorflow. International Journal of Innovative Research in Computer and Communication Engineering* 6, no. 7 (2018): 3888-3892.

[10] Azlah, Muhammad Azfar Firdaus, Lee Suan Chua, Fakhrul Razan Rahmad, Farah Izana Abdullah, and Sharifah Rafidah Wan Alwi. "Review on techniques for plant leaf classification and recognition." *Computers* 8, no. 4 (2019): 77. https://doi.org/10.3390/computers8040077

[11] Mikołajczyk, Agnieszka, and Michał Grochowski. "Data augmentation for improving deep learning in image classification problem." In *2018 international interdisciplinary PhD workshop (IIPhDW)*, pp. 117-122. IEEE, 2018. https://doi.org/10.1109/IIPHDW.2018.8388338

[12] Dar, Refaz Ahmad, Mohd Shahnawaz, and Parvaiz Hassan Qazi. "General overview of medicinal plants: A review." *The journal of phytopharmacology* 6, no. 6 (2017): 349-351. https://doi.org/10.31254/phyto.2017.6608

[13] Shakya, Arvind Kumar. "Medicinal plants: Future source of new drugs." *International journal of herbal medicine* 4, no. 4 (2016): 59-64.

[14] Zhang, Jinghua, Chen Li, Yimin Yin, Jiawei Zhang, and Marcin Grzegorzek. "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer." *Artificial Intelligence Review* 56, no. 2 (2023): 1013-1070. https://doi.org/10.1007/s10462-022-10192-7

[15] LeCun, Yann, Koray Kavukcuoglu, and Clément Farabet. "Convolutional networks and applications in vision." In *Proceedings of 2010 IEEE international symposium on circuits and systems*, pp. 253-256. IEEE, 2010. https://doi.org/10.1109/ISCAS.2010.5537907

[16] Din, Roshidi, Nuramalina Mohammad Na'in, Sunariya Utama, Muhaimen Hadi, and Alaa Jabbar Qasim Almaliki. "Innovative Machine Learning Applications in Non-Revenue Water Management: Challenges and Future Solution." *Semarak International Journal of Machine Learning* 1, no. 1 (2024): 1-10.

[17] Khalid, Fatimah, and Amirul Azuani Romle. "Herbal Plant Image Classification using Transfer Learning and Fine-Tuning Deep Learning Model." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 35, no. 1 (2024): 16-25. https://doi.org/10.37934/araset.34.3.1625

[18] Kader, Mohamed Mydin M. Abdul, Muhammad Naufal Mansor, Zol Bahri Razali, Wan Azani Mustafa, Ahmad Anas Nagoor Gunny, Samsul Setumin, Muhammad Khusairi Osman *et al.,* "Environmental Lighting towards Growth Effect Monitoring System of Plant Factory using ANN." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 43, no. 2 (2025): 167-177. https://doi.org/10.37934/araset.43.2.167177

[19] Jerome, Nirmal Jothi, Sivasankari Jothiraj, Saranya Kandasamy, Divya Ramachandran, Dineshkumar Selvaraj, and Poonguzhali Ilango. "An Effective approach for Plant Disease Detection Using Assessment-Based Convolutional Neural Networks (A-CNN)." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 31, no. 3 (2023): 155-172. https://doi.org/10.37934/araset.31.3.155172

[20] Pound, Michael P., Jonathan A. Atkinson, Alexandra J. Townsend, Michael H. Wilson, Marcus Griffiths, Aaron S. Jackson, Adrian Bulat *et al.,* "Deep machine learning provides state-of-the-art performance in image-based plant phenotyping." *Gigascience* 6, no. 10 (2017): gix083. https://doi.org/10.1093/gigascience/gix083