IDENTIFICATION OF MEDICINAL PLANTS USING DEEP TRANSFER LEARNING

BY

SAURAV BARAL

ID No: 213-15-4412

&

SHITHI RANI SAHA

ID No: 213-15-4280

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised by

Rahmatul Kabir Rasel Sarker
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH
JULY 2024

APPROVAL

This Project titled **"Identification of Medicinal Plants Using Deep Transfer Learning"**, submitted by **"Saurav Baral"**, ID No: **213-15-4412** & **"Shithi Rani Saha"**, ID No: **213-15-4280** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on July 16, 2024.

BOARD OF EXAMINERS

Narayan Ranjan Chakraborty (NRC)
Associate Professor & Associate Head — **Chairman**
Department of Computer Science and Engineering
Daffodil International University

Naznin Sultana (NS)
Assistant Professor — **Internal Examiner**
Department of Computer Science and Engineering
Daffodil International University

Raja Tariqul Hasan Tusher (THT)
Assistant Professor — **Internal Examiner**
Department of Computer Science and Engineering
Daffodil International University

Dr. Ahmed Wasif Reza (DWR)
Professor — **External Examiner**
Department of Computer Science and Engineering
East West University

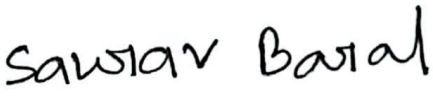
DECLARATION

We hereby declare that this project has been done by us under the supervision of **Rahmatul Kabir Rasel Sarker, Lecturer, and Department of CSE** at Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Rahmatul Kabir Rasel Sarker
Lecturer
Department of CSE
Daffodil International University

Submitted by:



Saurav Baral
Id Number: 213-15-4412
Department of CSE
Daffodil International University

Shithi Rani Saha
Id Number: 213-15-4280
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project successfully.

We really grateful and wish our profound our indebtedness to **Rahmatul Kabir Rasel Sarker, Lecturer, Department of CSE** Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "web application development" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori, Professor & Head**, **Department of CSE**, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

This thesis explores the identification of medicinal plants in Bangladesh through images of plant leaves using deep transfer learning. The primary aim of the study was to find the best and less time-consuming procedure, which could enable people to identify different types of medicinal plants by their leaves. To this end, we introduced and compared multiple deep learning models, which are MobileNet V2, Inception V3, ResNet50, VGG16 & VGG19. Based on their performance in the given tests, it was revealed that the MobileNet V2 had the best performance among all the four models. It obtained an average accuracy of 94.54% on the dataset, which is rather high. Regarding the performance characteristics, accuracy, recall rate, and F1-score are 95% each. Combining these results shows how the model is great at identifying medicinal plants from the leaves images. The research findings of this study have major managerial implications. Regarding this model, the following advantages can be provided: The possibilities for its application result in easier identification of medicinal plants by individuals, especially those from rural and other underserved regions, thereby increasing the use of traditional medicine. This tool may also be useful in an educational context since the nature of its approach is explanatory: students and researchers may use this tool to get more information on the diverse variety of medicinal plants found in Bangladesh. The model enhances environmental sustainability by supporting the accurate identification and that causes preservation of plant species. By reducing the risk of misidentification and over-harvesting, it plays a great role in the conservation of biological diversity. The research we presented highlights the potential of deep learning technologies, like MobileNet V2, in the prospects of effective relations between humans and the environment. Besides establishing the foundation for future research, this work also creates the roadmap for the creation of helpful and easy to use tools that may allow individuals to apply modern technologies to knowledge obtained from more conventional sources.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Plants, nature's silent masters of the world's surroundings, have enchained the imaginative element in man and sustained life from the beginning of the ages. From the antique zones of sequoias to the slender flowers that pave glades, plants cover a range of unbelievable different forms, sizes, and adaptations. They are the prime movers in terrestrial and marine conditions and therefore are in a fundamental role in shaping the planet's climate, environment and to some extent, the bio-diversity.

Medicinal plants have been needful to human existence and prosperity from early eras, being the main cure in distinctive maladies and disease. These plants, which frequently accommodate bioactive substances with recuperating attributes, have been a fundamental support of different conventional medication frameworks over the world. Beginning from the ancient dispatched of Ayurveda in India to the conventional Chinese restorative system and Traditional Chinese Medicine and natives of Local American tribes utilizing the inborn plants, the utilization of such plants remains boundless to space and time.

Restorative plants in Bangladesh secure a central view of its social status, diverse types of assorted life, and customary medicinal abilities. Having a rich culture that has always incorporated conventional medications, Bangladesh brags of a disparate class of restorative plants, which are conspicuous for treating multifaceted maladies. In this case, the discussion focuses on the origin, relevance, and contemporary importance of restorative plants in the context of Bangladesh considering their social, biological, and restorative roles.

Bangladesh, located in the mature delta region of the Ganges-Brahmaputra-Meghna Delta system, has proper climatic and geographical condition favorable for increasing the number and size of the important extent of plant species. The utilization of restorative plants can be dated a long time back when aboriginal populaces utilized the showcase nature to manage the wellbeing problems. This information has prolonged through generations and eras to form the groundwork of therapeutic

nails that subtly conform the current structured hones, thriving in portions of alfresco countryside even in the current generation.

The total number of plant species in the entire globe is roughly an astounding 380,000 [1]. Specifically, over 50,000 plants are reported to have medicinal uses within the world [2]. Medicinal plants are considered as those that are used to treat diseases. Medicinal plants used for Ayurveda and Unani medicines have been used since ancient times. The foremost Medicinal plants are used in China and India for which there are 11146 and 7500 kinds respectively [2]. In Bangladesh, 722 species of plants are medicinal plants according to the Bangladesh Agricultural Research Institute (BARI). Approximately 700 plants are utilized for medicinal purposes in Bangladesh [3].

Restorative plants in Bangladesh are perhaps one of the widely recognized plants in terms of social centrality deeply rooted in the country's history. It is worth underlining that the key role within the framework of disseminating information concerning restorative plants and their application belongs to conventional healers: the Kabirajes and Vaidyas. These healers, conventionally appreciated in their societies, will succeed stockpile of expounding received through proper training and practice. Theirs are honed with ceremonies, supplications, and chants, and, an integration of intelligence, muscle, and spirit in healing processes.

Besides, as social items restorative plants moreover contain colossal climatic importance in the present day's Bangladesh. The country consists of several biological systems which include mangrove forests, wetlands, tropical rainforests that consist a number of plant species having healing properties. However, there are significant threats associated with the uncontrolled deforestation, living space corruption and urbanization which are inimical to these valuable resources, and which also endanger plant diversity and conventional information networks.

In the most recent years, there has been a developing acknowledgment of the significance of moderating medicinal plants and advancing feasible gathering hones. Attempts have been made to index traditional knowledge, identify significant plant types, and establish conservation areas to save their taxa. Furthermore, efforts directed at increasing consciousness of communities situated near the productive use of medicinal plants and the need to preserve their natural habitats are gaining popularity.

The medicinal plants of Bangladesh are a great extent of species and they have special curing qualities. With reference to the neem tree (Azadirachta indica) persistent throughout the country because of its cleaning and anti-inflammatory properties, to the Tulsi (Ocimum tenuiflorum) which is adored for its medical and otherworldly importance among the plants, this forms an inseparable part of India's pharmacopeia. Other superb examples include aloe Vera, turmeric, ginger, and bitter melon, which are extensively used in conventional remedies for a multitude of ailments ranging from gastrointestinal upsets to skin disorders.

Utilization of medicinal plants in Bangladesh goes up beyond normal filled hones more than enhancing corrective associations and makes itself observed in upgraded medicinal associations too. Afterward, there has been an enhance fascination in home grown pharmaceutical in the later period because of the disputing side impacts of the administered drugs and a restores intension to herbal cures. Therefore, there has been emergences of home-grown items, supplements and make –up in the advertise, exploiting such a rich source of restorative plants as the nation has to offer.

In addition, research institutions, colleges, and pharmaceutical companies are practically engaged within the scientific study of medicinal plants, striving to sanction their preceding appliance, ascertain bioactive compounds, and design newly drugs. This integration of standard knowledge and cutting-edge rational schemes has the promise for the identification of new treatments for a broad spectrum of diseases; including cancer, diabetes, and cardiovascular diseases.

Emerging from the social relevance, it is also crucial to define the role of the medicinal plants in developing the conventional festivities, ceremonies, and events of Bangladesh. Several plants are valued for possessive spirit qualities and are mainly for use in religious rituals and cultural practices. For instance, basil (Tulsi) is not only held in esteem for its curative properties but also holds strict religious or cult values in the Hindu religion and is perhaps grown in backyards or temples. Altogether similar to the last one, another sacred tree of India is Ficus religiosa also called Bodhi tree which is central to Buddhism and the tree of enlightenment. These social associations add an extra dimension of meaning to restorative plants in the fabric of Bangladeshi society, it shows that their role is not just physical healing.

Medicinal plants are not just for healing. They also help the earth in big ways. They keep the soil together, clean the water, and give homes to animals. For instance, the Sundari tree, a type of

mangrove, gives us medicine but also fights shore wash away, offers a safe place for fish to grow, and helps people make a living through things like honey and wood. We see here how healing plants link with the health of our whole environment. This makes it clear we need to take care of these plants and use them wisely.

Today, in Bangladesh the additional emphasis is given to the traditional medical system and as such people have started to incorporate the use of the herbal treatment in their process of healing. Noticing an opportunity of extending possibilities of the modern medical therapies with the aid of the medicinal plants research, the tendency of introducing the supplements and formulations from the herbal remedies into the primary healthcare systems has recently emerged. on the basis of this approach phytotherapy or herbal medicine can be considered as integrative with a focus on the synergy of traditional knowledge and scientific innovations. But not only does it give patients an opportunity to get a broader range of treatments, traditional knowledge and ecosystem are saved as well.

Further, the overall economic uses of the medicinal plants can be a source of empowering this sector and opening up the idea of new and improved livelihoods in the rural zones of Bangladesh. The farming, processing and selling of the medicinal herbs and other related products could be source of income especially among the village people with traditional systematic knowledge about the plants. The government of Bangladesh can fully unlock economic value from the nation's tremendous botanical resources through supporting and encouraging medicinal plant business and processing in the chain.

In the present world, it is a well-known fact that Medicinal plants are an important source not only in the conventional system of Healing but also in the Advance system or modern health care system. Medicinal plants have been applied in health care for not less than two millennia. Nevertheless, identification and authentication of the medicinal plants are relatively complicated. The differentiation of the medicinal plants is very challenging event for persons, even for qualified botanist because the shape and colors in the plant species of medicinal plants are so much similar. That is why, necessity arises in the case of an automated approach for the identification of the medicinal plants. This method of identification shall be done using the medicinal plant Leaf images.

Leaves of the medicinal plants are different in some aspects that can be employed to confirm the legitimacy of the plant species. Vegetative 'physiognomy' includes parameters of leaves, and their size and shape, which can significantly differ between different plants. For instance, the Azadirachta indica (Neem) tree's leaves are long and pointed; on the other hand, the Streblus asper (Seura) tree has rigid leaves with an oval shape. The surface and outer coloring of the leaf also make up other features of the plant leaf. There are plants that have smooth and shiny foliage while there are those that have rough ones. The contours of the leaf together with the margins are other attributes that can be used in the categorization process.

1.2 Motivation

The research aims at simplifying the identification process on the plants for those who do not have prior information of the medicinal plants. Most of the conventional methods used in the identification of plants require one to have prior knowledge or even undergo through a certain training. It will be very almost impossible for the general public to fully understand.

In this context, it is crucial to find an efficient approach to share this knowledge with a wider audience, which can be possible by using new methods derived from machine learning and image processing. and enable ordinary people to recognize various medicinal plants with the help of images of the leaves.

It has the ability to transform the healthcare system because consumers of healthcare can easily get herbal storage and utilize them properly. In conclusion, this research aims at building a connection between current horticultural knowledge and Presentation that is easy to understand by everyone.

It encourages sustainable practices of collecting and using the medicinal plants, and the conservation activities because it enhances the connection of people with nature. This research aims at transforming the health care delivery systems while at the same time singing a sweet music of caring for our planet and its abundance for the later generations.

1.3 Research Question

The main research question of this study is:

- Is it possible to apply machine learning algorithm for image recognition to differentiate and classify medicinal plant species using Bangladeshi images accurately?
- The sub-questions are:
- What are the current approaches and some of the problems in identification of medicinal plants using deep transfer learning?
- How can the following difficulties be resolved: fluctuations in lighting, additional objects in the background, and photos' quality?
- How can we determine if the proposed model in this research is effective and efficient?

1.4 Expected Outcome

The expected output of the study is identification of medicinal plants in Bangladesh using plant leaf images would be the ideal method or model that would enable people to identify different plants of medicinal importance depending on their appearance captured in images.

It may entail the use of a predictive model maybe of machine learning type, based on a large database of medicinal plant leave images in order to a be able to classify the various species when presented with new leave images.

Additionally, the study may produce insights into the effectiveness of different image recognition techniques, the feasibility of using machine learning for medicinal plant identification in diverse ecological contexts like Bangladesh, and potential applications of such technology in biodiversity conservation, traditional medicine, and related fields.

Ultimately, the desired outcome is a practical tool or system that can assist researchers, conservationists, herbalists, and other stakeholders in accurately identifying medicinal plants in Bangladesh using readily available leaf image data.

1.5 Project Management and Finance

The Project management and finance for this thesis are as follows:

- The project would take almost one year, from July 2023 to June 2024.
- The domain and title selection takes one month.

- Literature review and finding an existing study takes one month.
- The dataset is collected over two months.
- The data preprocessing is required one month.
- The experimentation period takes five months.
- The proposed model is defined over one month.

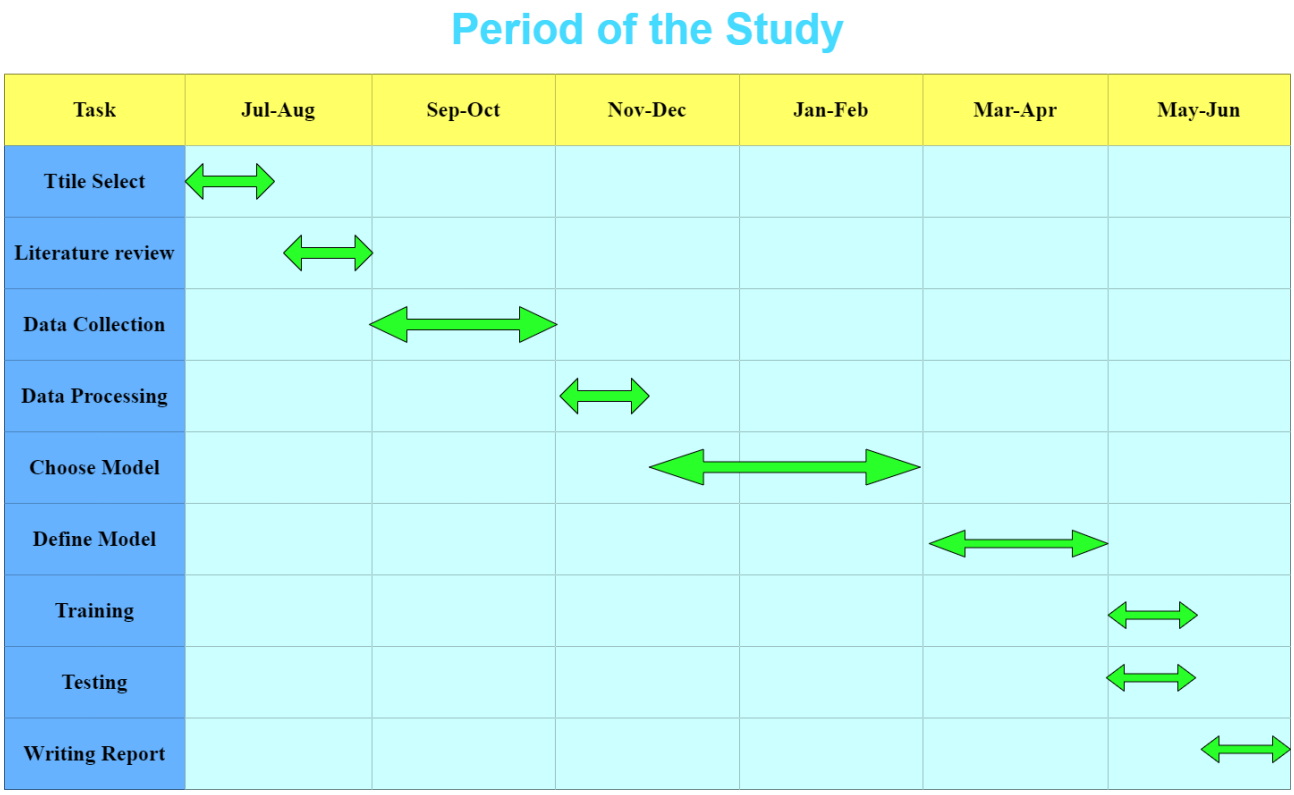


Figure1.1: Study life Cycle

CHAPTER 2

BACKGROUND

This chapter presents the theoretical background and literature search for the identification of medicinal plants by deep transfer learning.

The following sub-categories:

- Terminology: In this section, the reader is provided with the background information on the components of Computer Vision, including Machine Learning and Deep Learning.
- Related work: This section evaluates how previous methods and difficulties in determining the species and veracity of medicinal plants through image analyses.
- Comparative Analysis and Summary: In this context, this section carries out a summarization and comparison of the features, performances, and existing gaps of the method in the literature and limitations.
- Scope of the Problem: This section includes the research topic and questions of the study describes and elaborates the purpose and usefulness of the developed model.
- Challenges: This section describes the main problems and issues encountered in the course of the research and offers recommendations and approaches.

2.1 Terminologies

In this section, the background information and definitions that will elucidate the subject of the research work is provided. In this paper, we discuss Computer Vision and classify Machine Learning and Deep Learning as its sub-disciplines.

2.1.1 Computer Vision

Computer vision is considered a branch of artificial intelligence and computer science that relates to training computers to analyze and comprehend data received from the physical surroundings. But it concerns creating algorithms and methods through which the computers can analyze and understand images or videos. Computer vision applications are categorically grouped by tasks such as image recognition, object detection, image segmentation, image classification among others. Computer vision here imitates the visual perception of a human to a degree by allowing machines

to wake up in the physical environment, understand objects and patterns and make relevant decisions all by using visuals. Some of the domains where this technology is useful are, hospitals, self-driving cars, security cameras, smart glasses, and factories.

2.1.2 Machine learning

Artificial intelligence can be categorized into machine learning, which deals with the creation of models that allow computers to learn from the data to solve particular tasks without being taught. it is a process of creating programs that are able to find the patterns the predict or make conclusions about the data with the help of several iterations. Machine learning algorithms are trained on large datasets, where they identify patterns and relationships, allowing them to generalize and make accurate predictions on new, unseen data.

2.1.3 Deep learning

Deep learning is a subset of machine learning that uses artificial neural networks with multiple layers to learn data representations. The brain's structure inspires it and has succeeded in tasks like image and speech recognition, natural language processing, and medical diagnosis.

2.2 Related work

This section provides an overview of the previously used approaches and issues in the identification and authentication of medicinal plant species using images.

2.2.1 Traditional methods

The conventional approach to recognize medicinal plants in Bangladesh through images deals with the tribal and folk knowledge trade carried generation wise in the respective regions. Such approaches entail the process of recognizing images of medicinal plants by comparing them to their known morphological features. This paper also provides a general view of the previously used methodologies and they involve image preprocessing, feature extraction, and classification of medicinal plants in Bangladesh. Image pre-processing prepares the image for analysis by enhancing the image and its quality by narrowing down on the image features. Some of the

categories of techniques are noise reduction techniques, contrast enhancement techniques, techniques for resizing the images, normalization techniques, edge detection techniques, color correction techniques, registration techniques, and techniques for removing artifacts. In the application of these preprocessing techniques, it enhances the quality of the results obtained by the researchers. Attribute extraction is a kind of basic procedure in image analysis which is concerned with the identification of the content and/or structure of the data. It may comprise edges, textures, shapes, colors or other properties that may set off objects or sections in the scene, respectively within the image. Indeed, categorization is a basic task in supervised learning and pattern analysis where the main goal is to place the input data in one of the several pre-defined categories. This process often entails using a training set of images, in which an image belongs to precisely one labeled class. In the training phase, the model is able to identify patterns and features of images of classes that are being trained. Once it is trained, the task of the model is to determine the class label for new images that the algorithm has never encountered and extracts their features. Such methods can be supporting vector machines, decision trees, or deep learning neural networks which have the ability to learn and discover more features and relationships out of the data provided to them.

2.2.2 Recent methods

New approaches of the identification of the medicinal plants in Bangladesh by images are some of the developing and innovative methods and technologies that have been developed in the image analysis and machine learning fields. These methods rely on the progress that has been made in computer vision technologies and pattern and data recognition useful in plant identification by their images. The recent methods might comprise creating image recognition algorithms, deep learning models or CNN ensuring the images of medicinal plants from large datasets. These approaches help in the quick, precise and efficient identification of plant species which in turn helps the researchers, conservationists and plant users in Bangladesh.

2.2.3 Comparative Analysis

In the past decades, many researchers have proposed various works on plant identification. Some of them were based on low-level feature extraction such as color, texture, leaf vein extraction, etc.

The authors [4] proposed a method based on color, texture, and edge features for the identification and classification of Indian medicinal plants with 90% accuracy. The authors [5] proposed a method based on the texture and color features of digital leaf images for identifying Indonesian medicinal plants with 74.51% accuracy. The authors [6] proposed a method based on color, and edge features for the classification of medicinal plant leaf with 92% accuracy. The authors [7] proposed a method based on the canny edge detector and SVM classifier for Leaf Classification and Identification with 87% accuracy. The authors [8] proposed an automated system for the medicinal plant classification with 71.3% accuracy rate. The authors [9] introduced LeaVeS: A Leaf Recognition Application by Leaf Shape and Venation using Artificial Neural Networks that aim to classify a plant species by a leaf image alone. This method was able to achieve 95% accuracy in plant species classification. The authors [10] introduced Identification and Classification of Indian Medicinal Plants by combining color, texture, and edge features using SVM and radial basis exact fit neural network (RBENN). The method was able to achieve 74% and 80% accuracies for color, and edge texture features respectively. This method achieved 90% accuracy with combined color and texture features. The authors [11] introduced identification of medicinal plants using a hybrid transfer learning technique. This method was able to achieve 95.25% test accuracy. The authors [12] introduced Identification of Selected Medicinal Plants Leaves by extracting shape, color, and texture features from leaf images using an Artificial Neural Network (ANN) classifier. They got 94.4% accuracy from this method with a minimum of 8 input features. The authors [13] introduced Identification and Classification of tropical medicinal plants by extracting patterns from the leaf using SMO, Decision Tree (J48), Naïve Bayes (NB), kNN, Random Forest (RF), and Direct Ensemble Classifier for Imbalanced Multiclass Learning (DECIML) classifiers. They got 50%, 40%, 35%, 40%, 45%, and 65% accuracy respectively.

2.2.4 Scope of the problem

The issue here is that this research seeks to establish the most suitable method of analyzing data that is complex and dynamic.

The study will look into these aspects of the problem:

- Data Collection: The research will work with the data of diverse types and apply appropriate methods to extract, transform and load the data into a specific format.

- Data Processing: The specifics of using machine learning, statistical methods, and data visualization tools, as well the generation of new techniques for analyzing and interpreting big data are among the objectives of the research.
- Data Presentation: The study will identify specific ways of effective communication of analytical results and will develop sound approaches on how to present the findings of data analysis through dashboards, reports, and narratives.

2.2.5 Challenges

Various challenges in accomplishing its objectives:

- Data Quality: The research has to input raw data that are often fragmented, discordant, ambiguous and stale and the best ways of preprocessing, merging and updating must be used.
- Data Complexity: The research must adapt to others that stem from high-dimensional, heterogeneous and nonlinear characteristics of the data set and select an appropriate procedure that will reduce, transform and model the data set.
- Data dynamics: To contain the dataset, the research has to deal with a dynamic, streaming, and real-time data type, therefore, the methods need to address issues like monitoring, updating and forecasting of the data type.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter outlines the background to the research work and the procedure to be followed in the identification of medicinal plants using deep transfer learning. The following sub-sections:

3.1 Research Subject and Instrumentation: This sample gives an introduction of the study topic and the instruments that will be used for data collection and analysis.

3.2 Data Collection Procedure and Dataset Utilized: This section makes it clear the approach taken for data collection and more to the point the dataset utilized in the research process.

3.3 Evaluation Matrix and Analysis: This section provides details of the analysis that is used in relation to the hypothesis testing and response to questions of the study.

3.4 Proposed Methodology or Applied Mechanism: This section introduces the specific approach or practical method by which the paper aims at distinguishing mango leaf diseases through the convolutional transformer-based model and justifies why it is logical and efficient to use the method.

3.5 Implementation Requirements: This section need to state how the proposed methodology or the applied mechanism need to be implemented.

3.1 Research Subject and Instrumentation

- Research Subject: The study focus of this research is to identify medicinal plants through deep transfer learning. The major attention is paid to the selections of the plants which are used in the treatment in Bangladesh. These plants are said to be healthy plants and are usually used in African traditional systems of treatment. Correct identification of these plants is very vital for their proper use in treatment.
- Instrumentation: The major method applied in this study is deep transfer learning model. This model is for the classifying of plant species by the pictures of their leaves in a given image. For training and running this model, we used Google Colab establishing itself to a T4 GPU.

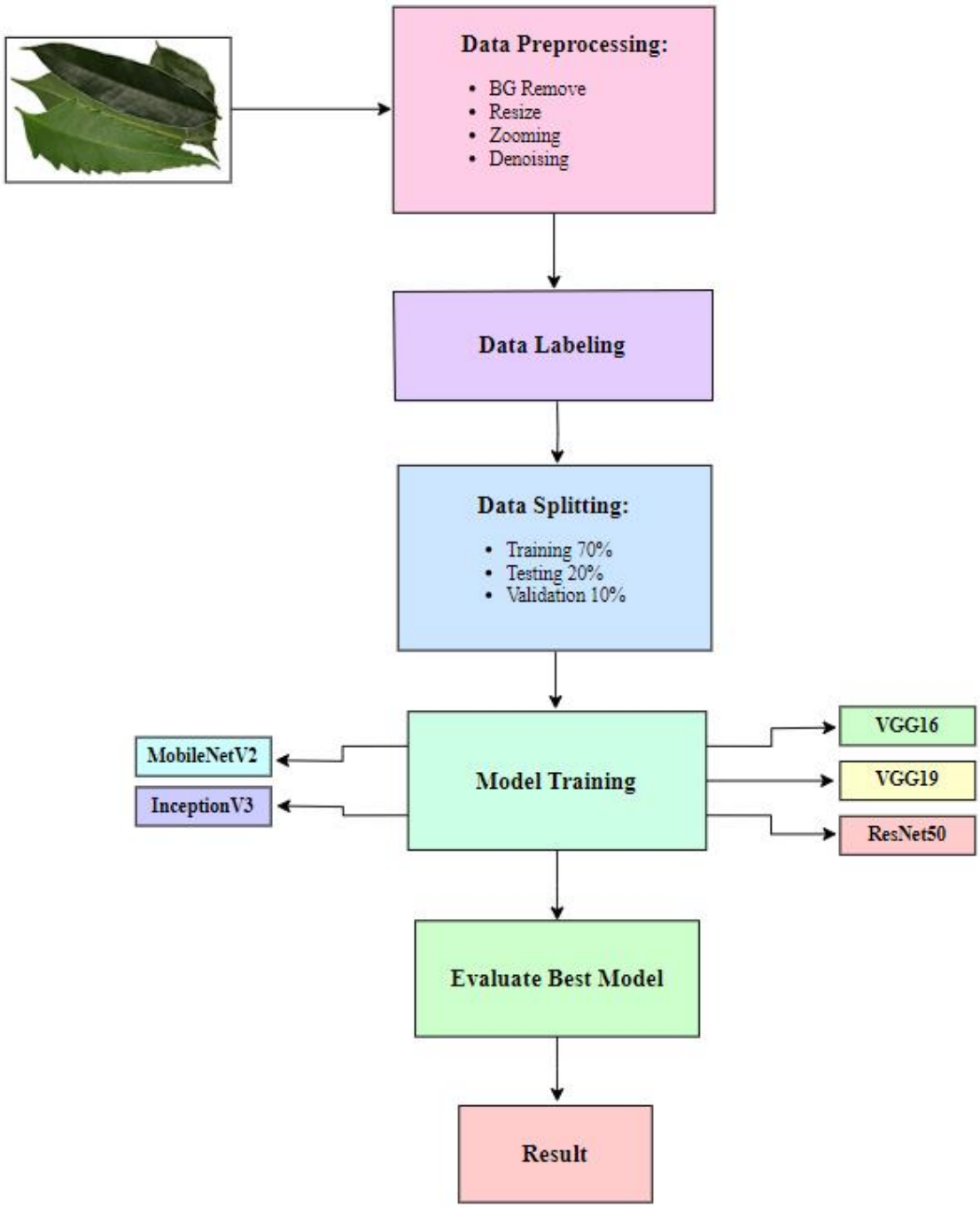


Figure3.1: Model Diagram

3.2 Data Collection Procedure and Dataset Utilized

Data Collection Procedure: The required set of plant images in the paper was assembled, including several hundreds of black and white images of plant leaves from different sources. These activities involved going to botanical gardens, forests, and other gardens where these plants can be located. All the images were appropriately commented with the name of the plant depicted. The comprehensive collection includes photographs of eleven medicinal plants: Ashok, Benjamina, Chitrashi, Dumur, Gutguitta, Jam, Kyaronja, Nagesshor, Neem, Sheura and Tulshi. Diversity in geographical variations of medicinal plants makes this paper appropriate for presenting a model for Bangladesh since the list focuses on medicinal plants. A total of 1265 images of single leaves of medicinal plants were captured using Canon EOS 250D DSLR Camera with an 18-55mm KIT Lens. That has 24. 1 MP resolution and DIGIC 8 image processor. The background of the images can sometimes introduce unwanted feature influences to the model during the training, and thus, to general features and, therefore, to exclude them as negative features during the training process, we use white background. Besides, each class's image was further augmented by applying the zooming and rotating to enlarge the size and scaling capability of the images and the final dataset included 2750 images of medicinal plant leaves from 11 classes where each of the classes contained 250 images. The dataset is divided into three subsets: A training set, a validation set and finally a testing set are used in this study. The training subset includes 70 percent of the data base containing 1925 images for the training of the proposed model. The validation subset is formed by 275 images (10 percent to the total data set) where the proposed model can be validated. The given testing subset is 20 percent of the overall set and has included 550 images for testing the given model.

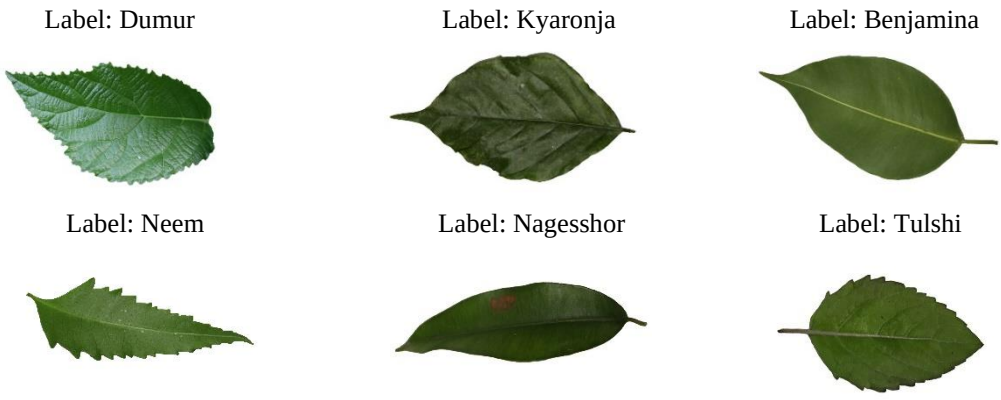


Figure3.2: Sample of Dataset

3.3 Evaluation Matrix and Analysis

- Evaluation Matrix: Regarding the performance of the proposed model, several evaluation metrics were leveraged when comparing the results.
- Accuracy: The proportion of plants recognized by the system among all the samples of plants.
- Precision: The number of plants that fall under each category in the evaluation of the models' accuracy, precision which is calculated as the ratio of correctly identified plants within the range of the total identified plants by the model.
- Recall: The accuracy of the final results against the total number of plants, counting for the model's uniqueness in identifying all instances it is used for.
- F1 Score: Average of precision and recall because you cannot optimize for both to get the optimal value.

3.4 Proposed Methodology or Applied Mechanism

Proposed Methodology: Our methodology involved several key steps:

- Data Preprocessing: Further, from the images that were extracted from the web, first, the images were rescaled to specific dimensions to make them more uniform. This comprised scaling, normalization, and data augmentation to diversify the dataset used in training the model.
- Model Selection: In this case, we employed a deep learning model which was previously trained and then trained the model on our own data set. The models are MobileNet V2, Inception v3, VGG16, VGG19, ResNet50. Using transfer learning, an already trained model is used without having to train it again on a different problem, and in this case, this enables us to reuse existing knowledge.
- Training the Model: The utilized model was pre-trained on our dataset; it was trained for several cycles to improve the outcomes and achieve the best result. Thus, we applied concepts such as cross-validation, which helped to avoid overlearning and ensure the stability of the final results.
- Validation and Testing: The performance of the model was the determined by validation and testing phases with other different datasets in order to fine-tune the model.

3.5 Implementation Requirements

The implementation is performed on Windows 11 on the Intel Core i3-1115G4 11th generation processor. It has a DDR4 RAM of 8GB. Processor Frequency is 3.00 GHz up to 4.10 GHz. Other requirements are:

- Software and frameworks: Python language is widely used not only as a tool for modelling but also for scripting. Keras are deep learning frameworks for building up and training deep learning models.
- Environment for Development: Interactivities such as development of models are done from tools like Jupyter Notebooks or Google Colab.
- Data Preprocessing: Based on the scaling, normalisation, and augmentation techniques, tools for preparing images are employed for making models to be less sensitive.
- Parameters for training: Make adjustments on the rate of training to improve the process of the model. Batch size discusses how to define an efficient batch size to improve the training speed. Epochs also define the number of training necessary to achieve the model's convergence. The function reveals the difference between the forecasted end and actual results and determines the appropriate parameters of the proposed model. The optimisers as for example Adam Optimizer should be used to maximize the efficiency of the weights' changes while conducting experiments.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup

We used deep learning algorithms to recognize medicinal plants, compared the algorithms' effectiveness and considered possible improvements. The tests were conducted through Colab with the utilization of Keras, Tensorflow, and Numpy. In the data visualization we used Matplotlib to make plots. The dataset includes 2750 photos. We used 70% of the images for training. For validation, we used 10% of the data and for testing 20% of the data set was used.

Table 4.1: Splitting the dataset for training, validating, and testing

Data Type	Data Point
Training	70%
Validation	10%
Test	20%

4.2 Experimental Results & Analysis

The performance of model was evaluated using some key metrics:

4.2.1 Loss

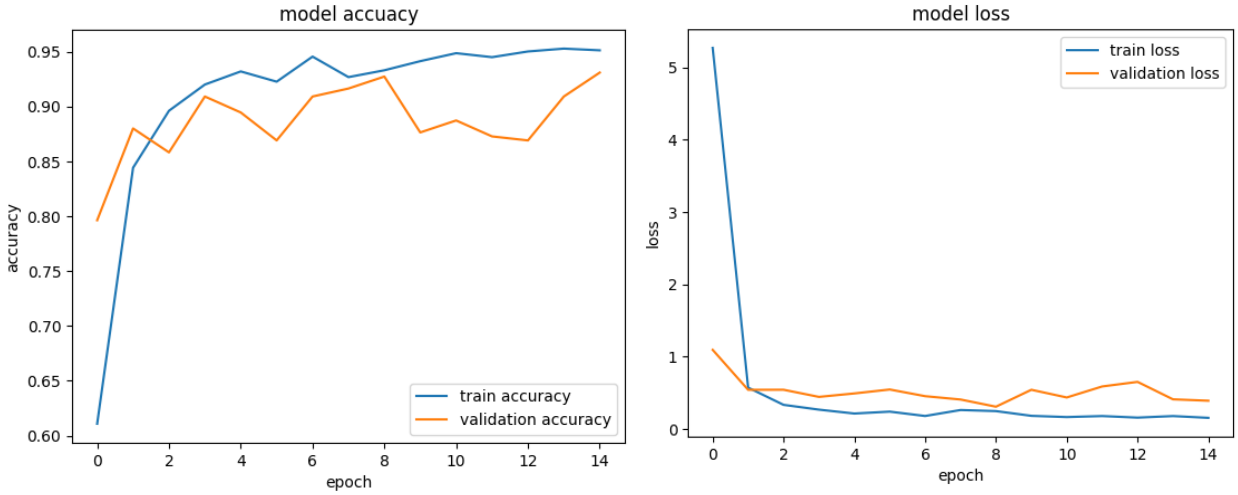
During the training and validation processes, categorical cross-entropy was paid attention to as the loss. A reduction loss meant that model is gradually improve in terms of recognize the medicinal plants.

4.2.2 Accuracy

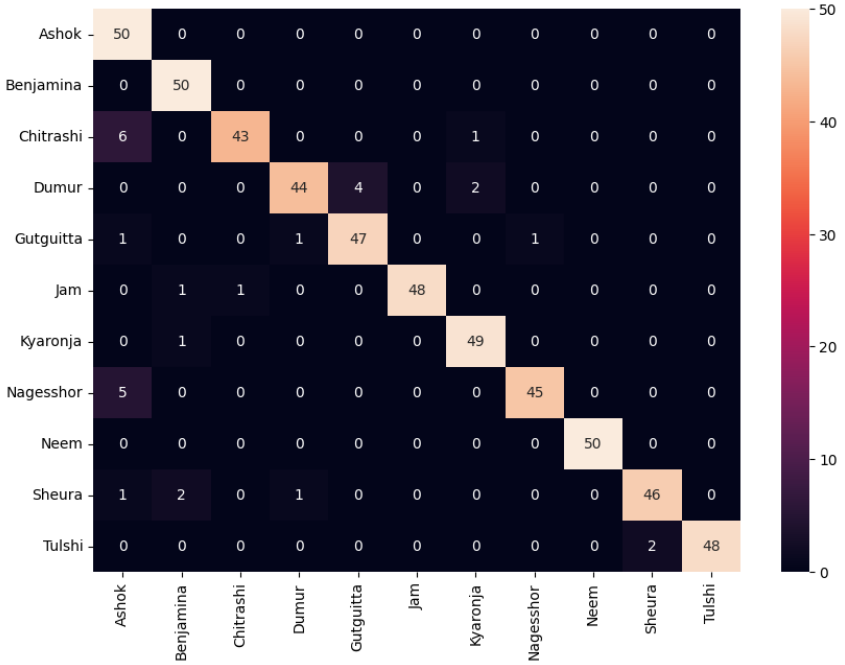
Accuracy of classification computed as the ratio of the number of responded samples to all the samples, offered an overall estimate of the efficiency of the model.

4.2.3 Confusion Matrix

The confusion matrix was used to analyze the behavior of the model to classify and predict the occurrences for each class; this comprised of true positive, true negative, false positive, and false negatives. This matrix gave an indication of some of the possible areas of emphasis and possible difficulties in identifying the medicinal plants.

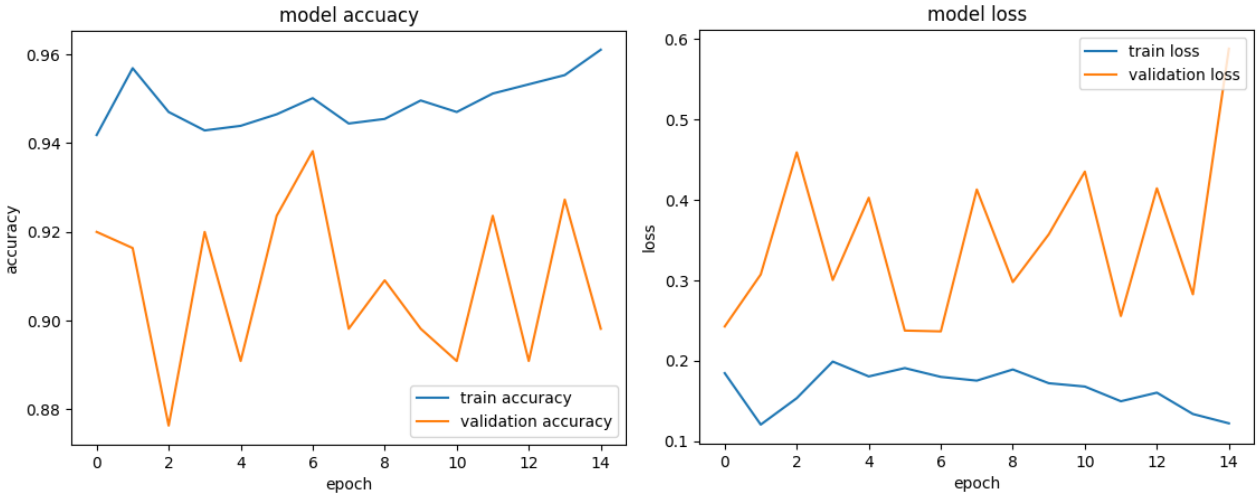


(a) (b)

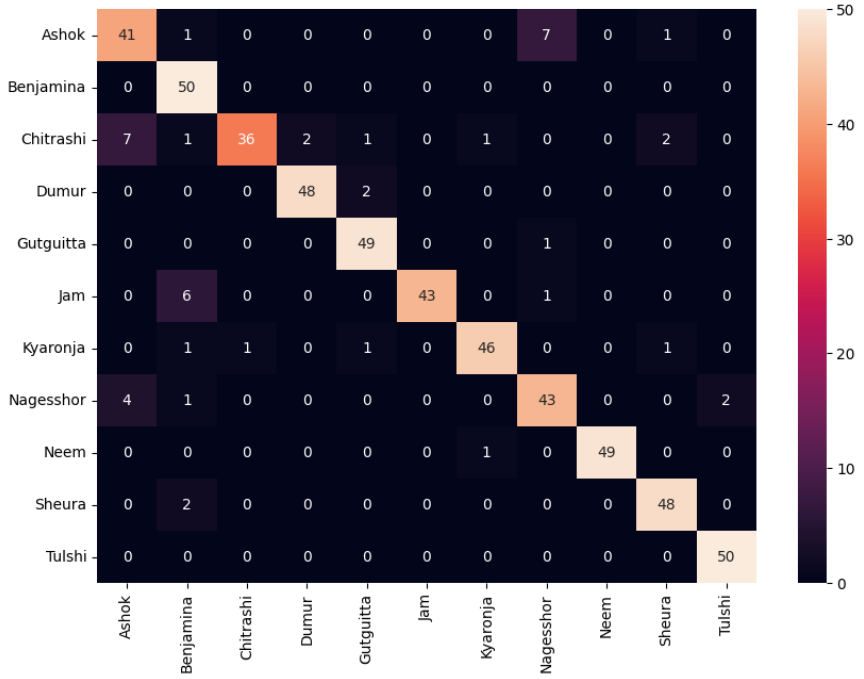


(c)

Figure 4.1: (a) Accuracy, (b) Loss, and (c) Confusion Matrix of MobileNet V2

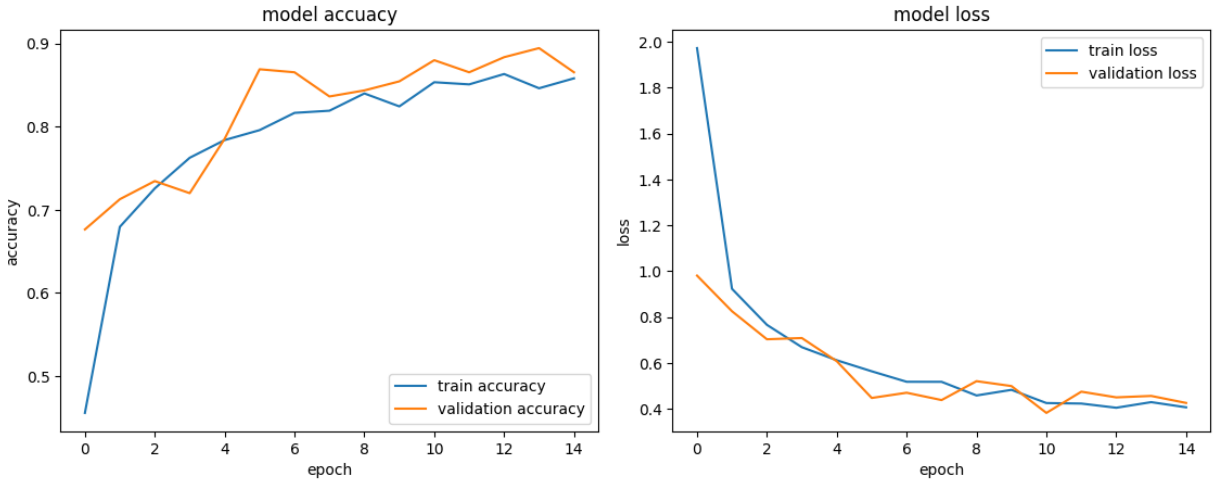


(a) (b)

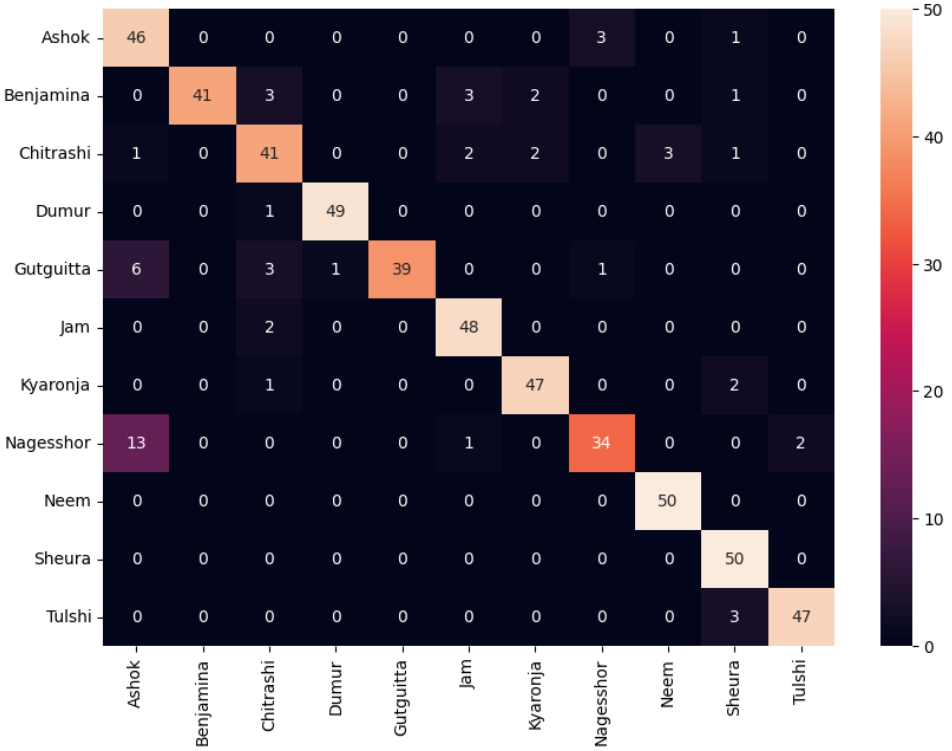


(c)

Figure 4.2: (a) Accuracy, (b) Loss, and (c) Confusion Matrix of Inception v3

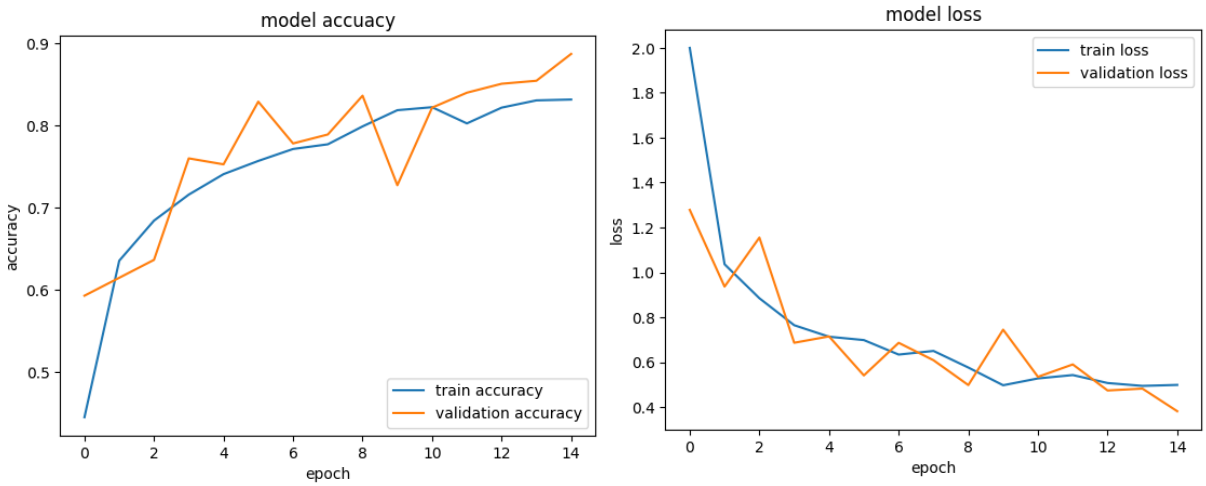


(a) (b)

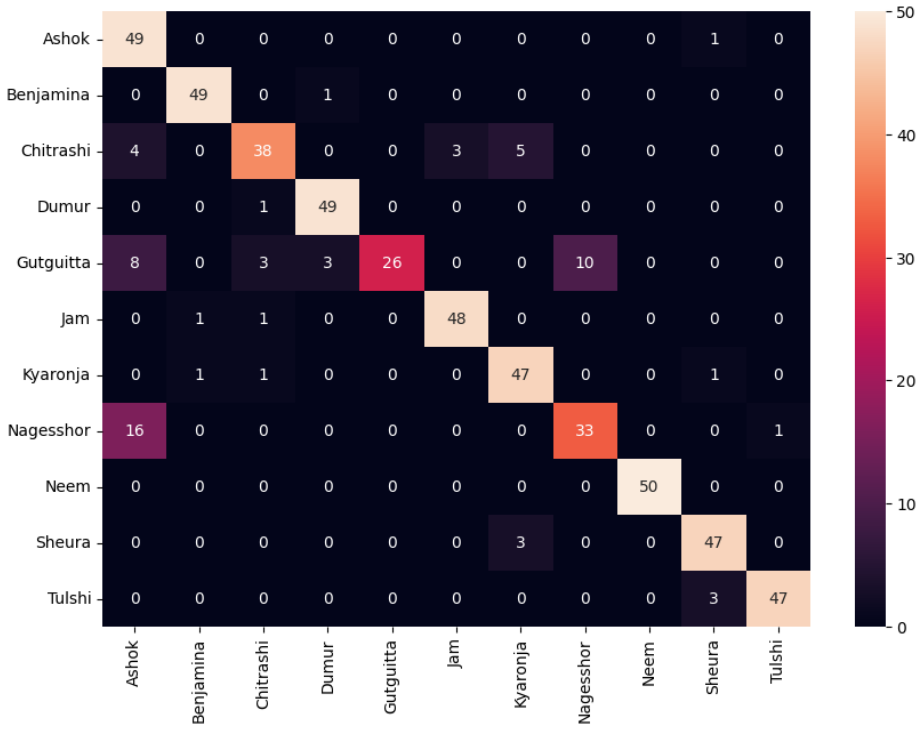


(c)

Figure 4.3: (a) Accuracy, (b) Loss, and (c) Confusion Matrix of VGG16



(a) (b)



(c)

Figure 4.4: (a) Accuracy, (b) Loss, and (c) Confusion Matrix of VGG19

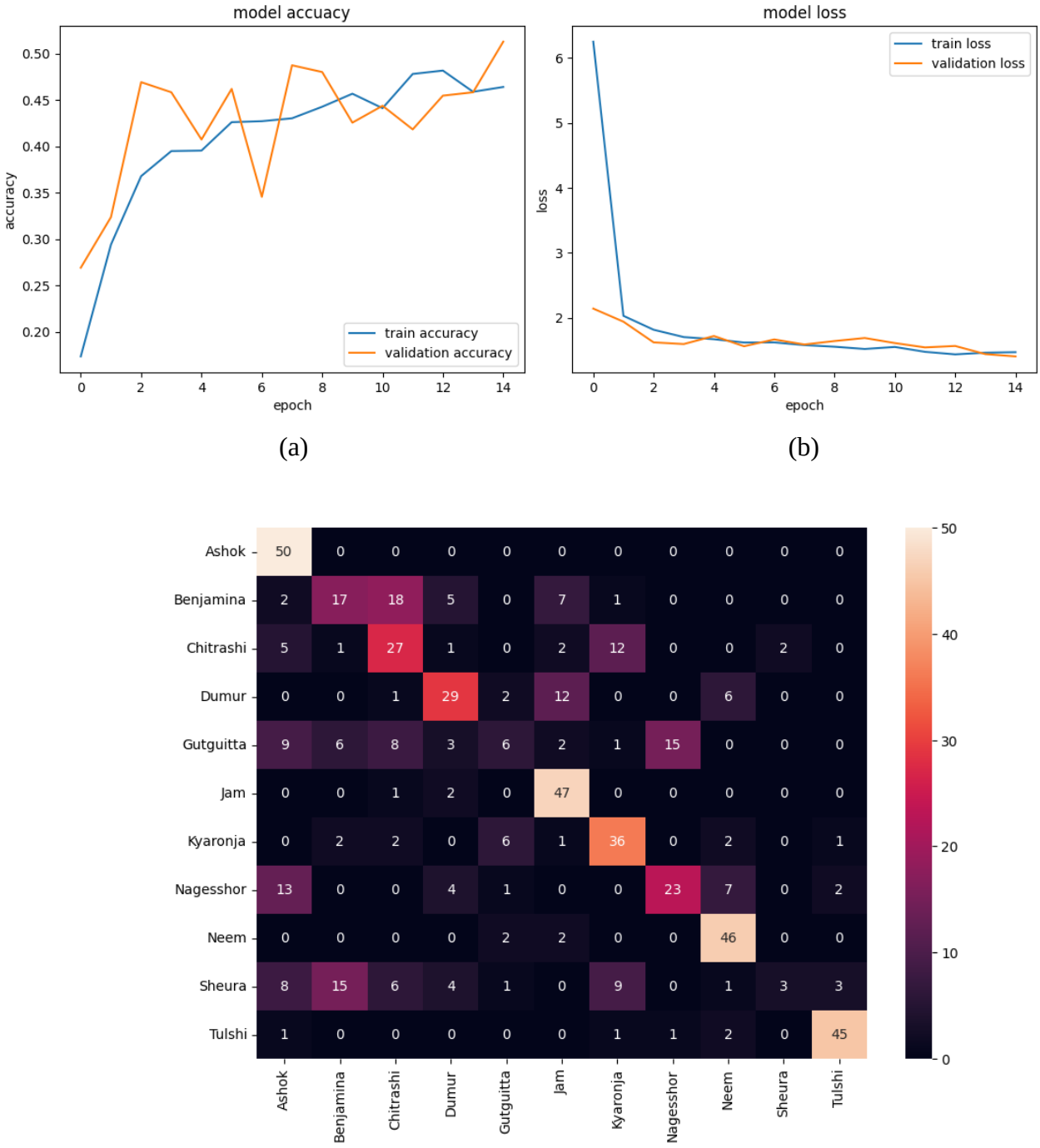


Figure 4.5: (a) Accuracy, (b) Loss, and (c) Confusion Matrix of ResNet50

4.3 Discussion

Machines are capable of identifying the medicinal plants correctly if the picture collection is well balanced as it is the prerequisite for training the deep learning models. To ensure the quality and the diversity of the images in the dataset completely new and relevant photos were collected from botanic gardens, forests and other gardens in Bangladesh. This, in return, has an impact on the ability of the model to classify as well as identify several characteristics of medicinal plants correctly.

The analysis of the gathered dataset was further used to recognize medicinal plants using several deep learning models such as MobileNet V2, Inception v3, VGG16, VGG19, and ResNet50. Out of these models, MobileNet is a regularization that provides faster detail and more efficient training compared to the other implementation and higher performance on the particular dataset. According to the results of the conducted investigation, it was identified that MobileNet V2 provided the highest average accuracy rate of 94.55%. While the Inception v3 achieves accuracy rate of 91.45% on average. ResNet50, VGG16, and VGG19 obtained 59.82%, 89.45%, and 87.81% accuracy respectively. It is crucial to admit that the influence of the number of parameters on the efficacy of a deep learning model depends on the specific assignment and data collection. Regarding the training, through the preformed experiments of having different architectures and parameters to be tested.

Precision: Precision measures how many of the plants identified by the model are correct. If the model says, from a specific plant 95 of 100 leaves are correct then the precision is 95%.

Recall: Recall measures how good the model is in finding all the correct plants. If the model successfully identified 95 of 100 leaves from a specific plant, then the recall is 95%.

F1-Score: The F1-Score combines precision and recall into one number, which gives it, a better view of how good or bad the model is. A high F1-Score indicates that the model is good in both prediction and recall. The F1-Score that our model got was 95% which is very good.

Support: Support refers to the number of actual instances of each plant class in the dataset. It depicts the number of samples the model had to learn from for the corresponding plant type. This is useful for evaluating the validity of the model's performance statistics.

Table 4.2: Proposed Model Summary and Performance by Each Label

Plant Name	Precision	Recall	Fl-score	Support
Ashok	0.79	1.00	0.88	50
Benjamina	0.93	1.00	0.96	50
Chitrashi	0.98	0.86	0.91	50
Dumur	0.96	0.88	0.92	50
Gutguitta	0.92	0.94	0.93	50
Jam	1.00	0.96	0.98	50
Kyaronja	0.94	0.98	0.96	50
Nagesshor	0.98	0.90	0.94	50
Neem	1.00	1.00	1.00	50
Sheura	0.96	0.92	0.94	50
Tulshi	1.00	0.96	0.98	50

Table 4.3: Inception V3 Model Summary and Performance by Each Label

Plant Name	Precision	Recall	Fl-score	Support
Ashok	0.79	0.82	0.80	50
Benjamina	0.81	1.00	0.89	50
Chitrashi	0.97	0.72	0.83	50
Dumur	0.96	0.96	0.96	50
Gutguitta	0.92	0.98	0.95	50

Jam	1.00	0.86	0.92	50
Kyaronja	0.96	0.92	0.94	50
Nagesshor	0.83	0.86	0.84	50
Neem	1.00	0.98	0.99	50
Sheura	0.92	0.96	0.94	50
Tulshi	0.96	1.00	0.98	50

CHAPTER 5

IMPACT AND SUSTAINABILITY

This chapter describes the impact of the research on the identification of medicinal plants in Bangladesh by using deep transfer learning in the social, environmental, and sustainability aspects of society.

It covers these sub-sections:

5.1 Impact on Society: In this section, it is illustrated how the research work can contribute positively to society and especially to the people of Bangladesh in the better identification and utilization of the medicinal plants. This research can contribute to health care, the development of traditional medicine, and assisting people in comprehending the natural ways to determine and treat illnesses using local plants.

5.2 Impact on Environment: This section shows how the research work would enable the protection of medicinal plants and their sustainable use. In its outcome, the research contributes to the diminution of synthetic medications' usage and encourages the use of natural treatments, which promote the maintenance of species diversification and genetic plant resources.

5.3 Ethical Aspects: This section also reveals and discusses the ethical considerations and issues concerning the research work; in this case, data privilege, safety and ownership of plant image data. It also takes into account some sociocultural effects of the technology and thus makes sure that the favorable outcomes of the study are shared fairly and gives regard to cultural practices and values.

5.4 Sustainability Plan: This section presents the sustainability plan of the research that is the practical continuation of the work, including the maintenance of the methodology and the deep transfer learning model as well as the scalability and accessibility of the developed solutions. The plan guarantees that the system will always be current, easily accessible by the public, and suitable to accommodate technological trends and botanological findings in the future.

5.1 Impact on Society

The research work on identification of medicinal plants using deep transfer learning can have a positive impact on public health and community well-being in various ways:

- Healthcare Accessibility: It has been observed that accessibility of modern healthcare services is still a major problem in many parts of Bangladesh and particularly in Rural and hard to reach areas. Since it will act as a guide to the relevant medicinal plants, people can take simple relief for widespread conditions through natural healers. This a way of minimizing the cocktail like use of expensive drugs and keep control of the heath costs.
- Empowerment through Knowledge: This implies that this technology will enable the empowering of individuals and communities by availing to them information on how they can use local resources. These facilities can empower women to become more independent and sustainable to some extent in cases of disease outbreaks or scarcities.
- Educational Value: This piece of technology can be adopted and included in the course work studies in schools and universities for the better experience of the students in botany, medicine and environmental studies. By serving an instrument for learning about plants, it will reintroduce natural sciences and give people more information about the role of medicinal plants.
- Cultural Preservation: The uses of medicinal plants are precondition of Bangladesh cultural sensation. In this way the technology known as plants also assists in preserving traditional knowledge that may otherwise not be passed on. This can help promote pride and Cim attribute continuity within those sub groups that benefit from it.
- Research Advancements: To the researchers, this model can act as a reference point when evaluating plant properties as well as their impact. It can contribute to a continuing search for the fresh facts concerning herbal remedy and it can open the ways to find new remedy or drug from natural resources.

5.2 Impact on Environment

The research work on identification of medicinal plants using deep transfer learning can have a positive impact on environment in various ways:

- Biodiversity Conservation: Some plants have specific uses in traditional medicine and when people over-exploit them, these certain species are eliminated. This way the technology can assist in point out more species of medicinal plants for harvesting and therefore spreading the pressure across a large number of species thus a positive view on the aspect of biodiversity.
- Sustainable Harvesting Practices: When the population of plants is supplemented through conservation and the right method of harvesting is taught to the public, the population should not reach its lower limit of sustainability. This can be quite relevant in the situation where specific plant species may be on the verge of being eradicated from the face of the earth.
- Reduction in Chemical Use: Through encourage usage of natural product man can reduce the usage of synthetic drugs and chemicals that pollute the environment during their preparation as well as when they are disposed. Natural remedies are environmentally friendly more so due to the fact that they are biodegradable products.
- Ecosystem Health: Efforts on the conservation also benefit the health of entire ecosystems with regard to the products derived from medicinal plants. Therefore, plants are vital structure components in the support of soil, the regulation of water cycle and being a home to many animals too. Hence, the protection of these medicinal plants will also help in protecting the balance as well as stability of ecosystems.

5.3 Ethical Aspects

The research work on identification of medicinal plants using deep transfer learning also involves some ethical aspects and challenges:

- Respect for Indigenous Knowledge: Another factor is the apprehension of indigenous cultural and historical rights and their knowledge on natural resources is appropriate. Efforts should be made to establish cooperation with these communities to reaffirm them individual and collective recognition for their intellectual property. This can include; Availing the gains accruing from the technology to the communities that have supplied the knowledge.

- Data Privacy and Consent: The model uses images of plants which may be originated from several sources. Any data that is gathered should be collected for the right reasons and with permission being sought. Some of the common principles include clarity to user and contributors on how data is gathered, saved, and utilized in the process.
- Accessibility and Equity: It would be best if technology were provided across the population and irrespective of the starting economic status of the people. There should be attempts to extend the use of this improvement towards the vulnerable groups of society that lack the privilege of access to quality health care services. This can involve releasing the tool in local languages and formats that are easily understandable to the users.
- Avoiding Exploitation: It means, there should be provision for preventing the technology from being utilised in a manner that is fraudulent for the commercial benefit of those who do not belong to the local community. Some general measures which should be set up to avoid the potential use of big data technologies for unethical purposes and to ensure that the key value orientations of big data applications remain the pursuit of the public good and sustainability.

5.4 Sustainability Plan

The research work on identification of medicinal plants using deep transfer learning also proposes a sustainability plan that involves several strategic actions:

- Continuous Model Improvement: To such an extent, the updates and refinements of the model can be frequent as the updated data take time to get improved. This presents a continuous need of working with technologist, botanist as well as local specialists who can feed the system with the most recent data on various plant species or their uses in medicine.
- Training and Education: Another measure that may be helpful for the society is to launch educational programs including holding corresponding workshops to teach people how to address the problems connected with the using of the technology. Consequently, training healthcare workers, educators, and other community leaders is one of the approaches to creating a core group of informed users who can share the innovation's benefits with others.
- Community Engagement: This implies that involving the locals in the development and implementation of the technology helps in achieving the set goals while considering their

way of thinking and doing things. Local monitoring and feedback can assist in modifying the technology to the specific context thus making the technology more relevant and efficient.

- Policy Integration: The technology can be adopted if implemented by the government services it can also work with other organization such as non-governmental organizations to improve health care and conservation policies in command. Government endorsement can help in the implementation of the sustainable features and guarantee that all the advantages of the technology are achieved.
- Partnerships and Collaboration: This is why it can be effective to establish cooperation with different stakeholders such as academic institutions, healthcare providers, and environmental organizations as this will form a supportive environment for the technology. Partnership can open rounded benefits and initiatives that can be achieved through collaboration as well as provide more rounded solutions for the problems of sustainability.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION, AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

In this research work, identification of medicinal plants was achieved through deep transfer learning, which can be regarded as a complex aspect of artificial intelligence. The first objective was to design a model for identification of medicinal plants from the image of the leaves which will help others specifically the people of Bangladesh to identify and use the plants. It is most relevant when traditional medicine is a significant component of health intervention and when the use of modem medical facilities may be a challenge.

During the course of the work, we gathered a rich dataset of plant leaf images for which the conv transformer-based model was trained. This model harnesses deep learning algorithms in a manner that enables it to classify the plants' features. It can be seen that the given model produced fairly good accuracy, and it could be considered as useful for both health care and schooling applications.

6.2 Conclusions

- Effectiveness of Deep Transfer Learning: To conclude, this work further affirmed that deep transfer learning in the confirmation of medicinal plants is very effective. In identifying plants from their leaf's pictures, the deep transfer learning method achieved high levels of accuracy.
- Accessibility and Usability: The developed model is very simple to use and any human who is not so technologically inclined will not face any problems in using it. This makes it an efficient tool for low reach and impoverished regions of the global society in that they can fully capitalize on the nutritional and medicinal value of plant products.
- Cultural and Environmental Benefits: Thus, facilitating the identification of medicinal plants the model contributes to the protection of traditional knowledge and the sustainable

use of natural resources. It also serves the objectives of conservation of species by discouraging over exploitation of species that are preferred by many people.

6.3 Recommendation

- Community Training Programs: For optimal results of this technology, it is suggested to integrate training programmers in the targeted communities. It is suggested that such programs should raise awareness of people on how best to adopt the model as well as how to harmoniously harvest the plants for medicinal use.
- Integration with Healthcare Services: Hence, health care providers especially those in the rural areas need to embrace this technology. This may assist in enhancing the implementation of the Primary Health Care (PHC) strategy and at the same time encourage the practice of integrating the Traditional Medicine systems with the modern systems.
- Educational Initiatives: This should be among the teaching tools listed in faculties of schools and universities. This can help in creating awareness of botany, and herbal medicine among the students as well as, passing on traditional knowledge.

6.4 Implication for Further Study

- Model Improvement: More often and in the future, research should be focused on the further development of the model together with the base that stores information about the different sorts of medicinal plants. This can be done by cooperating with botanists and other specialists of the region to gather more data as to flora.
- User Feedback and Adaptation: The collection of data from the users will be very crucial in the enhancing the implementation of the such technologies. To implement the proper changes towards increasing the utility of the model, their experiences of using the model and the difficulties can be identified.
- Cross-Regional Application: As for the limitations of this research, this is essential to emphasize that despite conducting analyses in Bangladesh, the proposed model could be applied further to other regions with considerable usage of medicinal plants. Research

could also look at whether similar structures have effectiveness in other cultures and other parts of the world or indeed other biomes.

- Integration with Other Technologies: In developing this model further, there is need to incorporate the other mobile technologies such as; mobile health apps, remote sensing to enhance efficiency and reach. It can open up more opportunities for the study of the environment and health and other related discipline.

REFERENCES

[1] "Plant," [Online]. Available: https://en.wikipedia.org/wiki/Plant.

[2] H. Y. H. L. Q. W. C. L. A. S. Shi-Lin Chen, "Conservation and sustainable use of medicinal plants: problems, progress, and prospects," Chen et al. Chin Med, 2016.

[3] D. H. U. Ahmed, "Prospects of medicinal plants in Bangladesh," 13 April 2022. [Online]. Available: https://www.tbsnews.net/thoughts/prospects-medicinal-plants-bangladesh-402402.

[4] U. S. S. P. Fawzi Mahomoodally, "Automatic Recognition of Medicinal Plants using Machine Learning Techniques," International Journal of Advanced Computer Science and Applications, pp. 166-175, 2017.

[5] N. K. S. W. Yeni Herdiyeni, "Mobile application for Indonesian medicinal plants identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram," International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp. 301-306, 2012.

[6] S. P. R. V. G. A Gopal, "Classification of selected medicinal plants leaf using image processing," International Conference on Machine Vision and Image Processing (MVIP), pp. 5-8, 2012.

[7] A. S. U. B. M. T. Adil Salman, "Leaf Classification and Identification using," International Conference on Inventive Systems and Control, pp. 1-4, 2017.

[8] M. I. H. Raisa Akter, "CNN-based Leaf Image Classification for Bangladeshi Medicinal Plant Recognition," Emerging Technology in Computing, Communication and Electronics (ETCCE), 2020.

[9] G. S. Azeil Louisse Codizar, "Plant Leaf Recognition by Venation and Shape Using Artificial Neural Networks," 7th International Conference on Information, Intelligence, Systems & Applications (IISA), 2016.

[10] S. S. N. A. G. Basavaraj S. Anami, "A Combined Color, Texture and Edge Features Based Approach for Identification and Classification of Indian Medicinal Plants," International Journal of Computer Applications (0975 – 8887), pp. 45-51, 2010.

[11] A. S. S. K. Sukanta Ghosh, "Identification of medicinal plant using hybrid transfer learning technique," Indonesian Journal of Electrical Engineering and Computer Science, pp. 1605-1615, 2023.

[12] A. G. R. Janani, "Identification of selected medicinal plant leaves using image features and ANN," International Conference on Advanced Electronic Systems (ICAES), pp. 238-242, 2013.

[13] T. K. G. R. A. Mohd Shamrie Sainin, "Malaysian Medicinal Plant Leaf Shape Identification and Classification," Knowledge Management International Conference (KMICe), 2014.

Saurav Baral_Identification of medicinal plants using deep transfer learning

ORIGINALITY REPORT

19%	17%	8%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
2	Submitted to Daffodil International University Student Paper	4%
3	www.researchgate.net Internet Source	1%
4	"Advances in Computing and Data Sciences", Springer Science and Business Media LLC, 2021 Publication	<1%
5	Azeil Louisse Codizar, Geoffrey Solano. "Plant leaf recognition by venation and shape using artificial neural networks", 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA), 2016 Publication	<1%
6	"Proceedings of the 2nd International Conference on Big Data, IoT and Machine	<1%