# AROGYA - AN INTELLIGENT AYURVEDIC HERB MANAGEMENT PLATFORM

Project ID: 2020-112

Project Proposal Report

N.J. Pathiranage – IT17129404

BSc (Hons) in Information Technology
Specializing in Software Engineering

Department of Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

February 2020

# AROGYA – AN INTELLIGENT AYURVEDIC HERB MANAGEMENT PLATFORM

Project ID: 2020-112

Project Proposal Report

(Proposal documentation in partial fulfilment of the requirement for the Degree of Bachelor of Science Special (honors) In Information Technology Specializing in Software Engineering)

BSc (Hons) in Information Technology

Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

February 2020

## DECLARATION OF THE CANDIDATE AND SUPERVISOR

I declare that this is my own work and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

| Name | Student ID | Signature |
|---|---|---|
| N.J. Pathiranage | IT17129404 | |
| | | |

The supervisor/s should certify the proposal report with the following declaration.

The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

……………………………..　　　　　　　　……………………

Signature of the supervisor:　　　　　　　　Date

……………………………..　　　　　　　　……………………

Signature of the co-supervisor:　　　　　　　　Date

## ABSTRACT

Ayurvedic medicines have a vital role in preserving physical and mental health of human beings. There are a considerable species of plants in our environment which will serve us as home remedies for almost all diseases and keeps us healthy when taken for daily consumption. Especially in Sri Lanka we have our own set of rare Ayurvedic herbs which have been utilized by generations as medicinal treatments for a variety of diseases. Identification and classification of medicinal plants are essential for better treatment. But most of us are unable to identify these plants due to lack of knowledge. Lack of experts in this field makes proper identification and classification of medicinal plants a tedious task. Hence, a fully automated system for medicinal plant classification is highly desirable. Automatic classification of trees from leaves is a popular computer vision/machine learning task and has important applications. With this approach, there are existing applications which can identify plants with low prediction accuracies. However, these applications are based on foreign plant data sets that do not include valuable herbs and shrubs with medicinal qualities. Hence this research proposes a mobile application which can identify a group of Ayurvedic herbs in Sri Lanka using Convolution Neural Network models (CNN) with Transfer learning. The major objective of the study is to analyze the noisy image set using deep neural network architectures based on transfer learning, choose the best architecture and create a deep learning model that can be applied effectively for an application. The methodology includes gathering data, and transfer learning based on deep Convolutional Neural Networks used on noisy image set for processing them using TensorFlow in a local computer. Images will be retrained on the available neural network architectures, fine-tuned from pre-trained weights and then the best technique will be selected. The selected algorithm will be fine-tuned using data augmentation techniques on the labeled dataset and hyper-parameter tuning. Conclusively, the outcome of this study will be used by locals in identifying herbal plants accurately.



**Keywords: Ayurvedic Leaf Classification, Computer vision, Convolutional Neural Network, deep learning, transfer learning**

# TABLE OF CONTENTS

## LIST OF ABBREVIATIONS

ANN-Artificial Neural Network

CNN-Convolutional Neural Network

MSF-CNN-Multi-Scale fusion Convolutional Neural Network

DNN-Deep Neural Network

WHO-World Health Organization

HOG- Histograms of Oriented Gradients

LBP- Local Binary Patterns

SVM- Support Vector Machines

ELM- Extreme Learning Machines

PCA- Extreme Learning Machines

KNN- K-Nearest Neighbor

PNN- Probabilistic Neural Network

CBIR- Content-Based Image Retrieval

## LIST OF FIGURES

## LIST OF TABLES

# 1. INTRODUCTION

Ayurveda is an ancient medicinal system evolved in India around thousands of years ago, still followed by many people as it is purely natural and has no side effects [2]. It is very relevant from ancient to this most modern time because of its power to cure chronic diseases [2]. The parts like leaf, flower, root, bark and fruit are mainly used in the preparation of medicines in Ayurveda [2]. According to the World Health Organization (WHO) in 2009, 80% of the people around the world still rely on ayurvedic botanical drugs or medicine. They have significant contributions towards human lives and play a predominant role in the well-being of the global population.

According to Ayurveda every plant on earth has some medicinal value, so it is important to protect the plant and identify its medicinal values [30]. Studies have proved that consuming so much of allopathic medicines may lead to side effects as it carries out many chemical reactions within the body [30]. A general fact about Allopathy is that once it is taken, it requires taking another medicine to cure the side effects which has happened due to the previous medicine [30]. In general, process of consuming medicines will not end. Allopathic treatments are meant to treat the Symptoms of a disease whereas Ayurveda treats the root of the disease [30].

One of the major advantages of Ayurveda is that it does not have any side effects as it is purely natural, that is relevant in this most modern time as new diseases evolve due to changed lifestyle and changed diet [4]. Also, majority prefer Ayurveda medicine over Western medicine due to some versatile factors such as Completely natural, Massage treatment, Positive influence on mental health, Fights inflammation, adds to weight loss, Improves heart health, Healthy detox, Individually prescribed methods, etc. [4]. So, it is important for every human being to return to Ayurveda. Almost all general diseases can be cured through Ayurveda using the shrubs and herbs that are around us. Ayurveda also brings lots of foreign money to the country since many foreign countries are inclining towards it. Nowadays there is a potential class of customers who prefer ayurvedic medicines than other medicines. So, the knowledge about medicinal plants carried by generations must be preserved and protected.

Computer vision, pattern recognition, and image processing technologies provide promising results for identification and classification of medicinal plants. Identifying a medicinal plant with required medicinal values is one of the major challenging tasks [20]. It plays a crucial role in the preparation of ayurvedic medicines. Also, proper classification of medicinal plants is important for agronomist, botanist, ayurvedic medicinal practitioners, forest department officials and those who are involved in the preparation of ayurvedic medicines [20]. But lack of expert taxonomists is a major issue in this area. Even though herbal medicine has no side effects, treatment using a wrongly identified medicinal plant may claim the life of a patient.

Computer vision and image processing methods can bridge the gap between the lack of expert taxonomists and potential requirement in identification and classification of the medicinal plants. Features of leaf, flower, trunk, and branch are normally being used by the taxonomists for classification purposes [33]. According to many researches carried out regarding this problem [5-10], there are existing applications which can identify plants with low prediction accuracies. However, these applications are based on foreign plant data sets that do not include valuable herbs and shrubs with medicinal qualities. Additionally, in most of those researches, shape, color, and textures have been considered as the important spatial and morphological features to classify the plants [32]– [34]. Additionally, since the leaves are available on all seasons, most of the researchers have decided as it is the best choice for plant classification. But the color feature is not a judicious feature for classification since it varies on seasons and different stages of the same leaf will have a different color. Taxonomists use the variations in leaf characteristics as a tool for the classification of medicinal plants. Many research works have been done in this area. But the high rate of inter-class similarity in shape, color and texture features make this problem still a challenging and unsolved one. Hence, a fully automated system to correctly classify the local medicinal plants is inevitable at this point of time.

As a solution for all the mentioned research problems, this research work proposes a mobile application which can identify Ayurvedic plants using Convolution Neural Network models (CNN) with Transfer learning. Several CNN architectures will be applied in this study experimentally to obtain the best accurate model. A newly prepared and annotated Ayurvedic plant dataset from Sri Lanka will be trained on CNN from scratch. However, training a model from scratch is too costly. To overcome this challenge transfer learning will be applied in the Ayurvedic dataset. This model will be then used to build a mobile application in the Android platform with TensorFlow-Lite which can identify Ayurvedic plants using the built-in camera module.

## 1.1 Background & Literature Survey

An overall idea about the importance of Ayurveda in Sri Lanka, current problems faced by people in this domain, apparent tasks that should be there in any automated identification and classification as well as the limitations of previous researches carried out regarding automated leaf classification worldwide have been given in the introduction section. This section brings several works and important attention in the focus of this research which is plant species identification and classification.

As mentioned in the introduction section, researchers have tried many methodologies to extract the features and identify the plant species automatically. According to previous work done, there are existing applications which can identify plants with low prediction accuracies. However, these applications are based on foreign plant data sets that do not include valuable herbs and shrubs with medicinal qualities. Additionally, most of these methods make use of the combination of many parameters like color, shape, texture features etc.

In 2019, the research paper "Oak Leaf Classification: An Analysis of Features and Classifiers" [14] has presented a new oak leaf dataset and preliminary results for classification of 8 types of oak trees. The novelties include comparative analysis of a small set of hand-crafted geometric features and popularly used high-dimensional appearance features, such as Local Binary Patterns (LBP) and Histograms of Oriented Gradients (HOG). They have further compared commonly used Support Vector Machines (SVM) classifier with a recently popular, fast and robust learner called Extreme Learning Machines (ELM). Results indicate that a small set of geometric features reach an accuracy of 75%, while high dimensional appearance features can boost the performance up to92%.

In 2019, in the study of "Combination of Deep Features and KNN Algorithm for Classification of Leaf-Based Plant Species" [15], they have proposed an approach based on the combination of deep architectures. Deep features were extracted from the plant leaves using the fc6 layer of the previously trained AlexNet and VGG16 models. Then, the reduction of the number of deep features by using the Principal Component Analysis (PCA) method was done quickly and the best distinguishing features were obtained. Finally, the classification performances were calculated using the K-Nearest Neighbor (KNN) method. Flavia and Swedish plant leaf data sets were used to test the proposed system. According to the experimental results, the accuracy scores for Flavia and Swedish data sets was obtained as 99.42% and 99.64%, respectively.

In 2019, in the study of “AyurLeaf: A Deep Learning Approach for Classification of Medicinal Plants”[20], they have proposed AyurLeaf, a Deep Learning based Convolutional Neural Network (CNN) model, to classify medicinal plants using leaf features such as shape, size, color, texture etc. This research work also has proposed a standard dataset for medicinal plants, commonly seen in various regions of Kerala, the state on southwestern coast of India. The proposed dataset contains leaf samples from 40 medicinal plants. A deep neural network inspired from Alexnet is utilized for the efficient feature extraction from the dataset. Finally, the classification is performed using Softmax and SVM classifiers. Their model, upon 5-cross validation, achieved a classification accuracy of 96.76% on AyurLeaf dataset.

In 2019, in the study of “Fine-Grained Plant Identification using wide and deep learning model” [21], they have proposed a model to address the fine-grained plant image classification task by using the wide and deep learning framework which combines a linear model and a deep learning model. Proposed method sums the result of the wide and deep learning model using a logistic function so that discrete features can be considered simultaneously with continuous image content. Their works have used metadata such as the date of flowering and locational information for the wide model. Their experiment shows that the proposed method gives better performance than a baseline method.

In 2018, the study of “A Multiscale Fusion Convolutional Neural Network for Plant Leaf Recognition” [16], a multiscale fusion convolutional neural network (MSF-CNN) is proposed for plant leaf recognition at multiple scales. First, an input image is down-sampled into multiples low resolution im- ages with a list of bilinear interpolation operations. Then, these input images with different scales are step-by-step fed into the MSF-CNN architecture to learn discriminative features at different depths. At this stage, the feature fusion between two different scales is realized by a concatenation operation, which concatenates feature maps learned on different scale images from a channel view. Along with the depth of the MSF-CNN, multiscale images are progressively handled, and the corresponding features are fused. Third, the last layer of the MSF-CNN aggregates all discriminative information to obtain the final feature for predicting the plant species of the input image. Experiments show the proposed MSF-CNN method is superior to multiple state-of-the art plant leaf recognition methods on the MalayaKew Leaf dataset and the Leaf Snap Plant Leaf dataset.

In 2018, the study of “Improving Plant Recognition using Hybrid features from Connectionist and Knowledge-Based Approaches” [22] has proposed architecture that combined knowledge-based approach to improve the accuracy of plant recognition. Towards this, hybrid features are constructed by merging three types of knowledge-based features; morphological feature, texture feature and color feature with convolutional neural network extracted features. Their architecture consists of three main stages which are data pre-processing, feature extraction and classification. Before features are extracted, images will be resized and augmented in the pre-processing stage. To classify the species of leaf, they consider decision tree and artificial neural network as a classifier. They have experimented on two datasets: Flavia and Swedish dataset. The experimental result shows that the proposed architecture can predict unseen images correctly more than existing models.

In 2017, the study of “Plant Classification using Convolutional Neural Networks” [23] has proposed a Convolutional Neural Network (CNN) architecture to classify the type of plants from the image sequences collected from smart agro-stations. First challenges introduced by illumination changes and deblurring are eliminated with some preprocessing steps. Following the preprocessing step, Convolutional Neural Network architecture is employed to extract the features of images. The construction of the CNN architecture and the depth of CNN are crucial points that should be emphasized since they affect the recognition capability of the architecture of neural networks. In order to evaluate the performance of the approach proposed in this paper, the results obtained through CNN model are compared with those obtained by employing SVM classifier with different kernels, as well as feature descriptors such as LBP and GIST. The performance of the approach is tested on dataset collected through a government supported project, TARBIL, for which over 1200 agro-stations are placed throughout Turkey. The experimental results on TARBIL dataset confirm that the proposed method is quite effective.

In 2016, the study of “Ayurvedic Plant Species Recognition using Statistical Parameters on Leaf Images” [30], has proposed a simple and efficient methodology for Ayurvedic plant classification using digital image processing and machine vision technology. The three major phases in proposed methodology are pre-processing, feature extraction and classification. Pre-processing is done in order to highlight the relevant features to be used in the proposed methodology as well as to reduce unwanted noise from the input image, which reduces the chance of getting optimal feature values. In feature extraction phase, different morphologic features such as mean, standard deviation, convex hull ratio, isoperimetric quotient, eccentricity and entropy are extracted from the pre-processed leaf image. In the third phase, a new approach to classify ayurvedic plant species is adopted to recognize plant species by calculating the leaf factor of the input leaf using the extracted feature values and it is compared with the trained values that are stored in the database. An accuracy of 93.75% is obtained for the proposed methodology.

In 2015, the study of “Recognition of Whole and Deformed Plant Leaves using Statistical Shape Features and Neuro-Fuzzy Classifier” [24] has proposed a methodology for recognition of plant species by using a set of statistical features obtained from digital leaf images. As the features are sensitive to geometric transformations of the leaf image, a preprocessing step is initially performed to make the features invariant to transformations like translation, rotation and scaling. Images are classified to 32 pre-defined classes using a Neuro fuzzy classifier. Comparisons are also done with Neural Network and k-Nearest Neighbor classifiers. Recognizing the fact that leaves are fragile and prone to deformations due to various environmental and biological factors, the basic technique is subsequently extended to address recognition of leaves with small deformations. Experimentations using 640 leaf images varying in shape, size, orientations and deformations demonstrate that the technique produces acceptable recognition rates.

In 2015, the study of “A Convolutional Neural Network for Leaves Recognition Using Data Augmentation” [25], a seven-layer ConvNet using data augmentation is proposed for leaves recognition. First, they implement multiform transformations (e.g., rotation and translation etc.) to enlarge the dataset without changing their labels. This novel technique recently makes tremendous contribution to the performance of ConvNets as it is able to reduce the over-fitting degree and enhance the generalization ability of the ConvNet. Moreover, in order to get the shapes of leaves, they sharpen all the images with a random parameter. This method is similar to the edge detection, which has been proved useful in the image classification. Then have trained a deep convolutional neural network to classify the augmented leaves data with three groups of test set and finally find that the method is quite feasible and effective. The accuracy achieved by their algorithm outperforms other methods for supervised learning on the popular leaf dataset Flavia.

In 2015, the study of “DEEP-PLANT: PLANT IDENTIFICATION WITH CONVOLUTIONAL NEURAL NETWORKS” [26] studies convolutional neural networks (CNN) to learn unsupervised feature representations for 44 different plant species, collected at the Royal Botanic Gardens, Kew, England. To gain intuition on the chosen features from the CNN model (opposed to a ’black box’ solution), a visualization technique based on the deconvolutional networks (DN) is utilized. It is found that venations of different order have been chosen to uniquely represent each of the plant species. Experimental results using these CNN features with different classifiers show consistency and superiority compared to the state-of-the art solutions which rely on hand-crafted features.

In 2014, the study of “Ayurvedic leaf recognition for Plant Classification” [31], the performance of different features extraction methods are compared, different combinations of features and a number of classifiers applied for leaf identification process are also discussed.

In 2012, the study of “An Efficient Leaf Recognition Algorithm for Plant Classification Using Support Vector Machine” [27] uses an efficient machine learning approach for the classification purpose. This proposed approach consists of three phases such as preprocessing, feature extraction and classification. The preprocessing phase involves a typical image processing steps such as transforming to gray scale and boundary enhancement. The feature extraction phase derives the common DMF from five fundamental features. The main contribution of this approach is the Support Vector Machine (SVM) classification for efficient leaf recognition. 12 leaf features which are extracted and orthogonalized into 5 principal variables are given as input vector to the SVM. Classifier tested with flavia dataset and a real dataset and compared with k-NN approach, the proposed approach produces very high accuracy and takes very less execution time.

In 2012, the study of “Classification of Selected Medicinal Plants Leaf Using Image Processing” [32] has aimed at implementing a system using image processing with images of the plant leaves as a basis of classification. The software returns the closest match to the query. The proposed algorithm is implemented, and the efficiency of the system is found by testing it on 10 different plant species. The software is trained with 100 (10 number of each plant species) leaves and tested with different plant species) leaves. The efficiency of the implementation of the proposed algorithms is found to be 92%.

In 2010, the study of “Plant Leaf Classification Using Plant Leaves Based on Rough Set” [28] a method of plant leaf classification is proposed based on the neighborhood rough set. They mainly show that this is applicable to plant leaf classification. Experimental results on plant leaf image database demonstrate that the proposed method is effective and feasible for leaf classification.

In 2008, the study of “A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network” [29] they employ Probabilistic Neural Network (PNN) with image and data processing techniques to implement a general-purpose automated leaf recognition for plant classification. 12 leaf features are extracted and orthogonalize into 5 principal variables which consist the input vector of the PNN. The PNN is trained by 1800 leaves to classify 32 kinds of plants with an accuracy greater than 90%. Compared with other approaches, their algorithm is an accurate artificial intelligence approach which is fast in execution and easy in implementation.

### 1.2 Research gap

In computer vision, despite many efforts [8–13], plant identification is still considered a challenging and unsolved problem. It is very challenging since rich plant leaf morphological variations, such as sizes, textures, shapes, venation, and so on. Most existing plant leaf methods typically normalize all plant leaf images to the same size and recognize them at one scale, resulting in unsatisfactory performances [16].

In Sri Lanka we have our own set of rare Ayurvedic herbs. But most of us are unable to identify these plants due to lack of knowledge. Dissemination of knowledge regarding herbal plants is restricted mainly to very limited group of people and is passed down from generation to generation who practice traditional medicine.

So, recognizing these endemic herbal plants is a challenging problem in the fields of Ayurvedic medicine, computer vision, and machine learning, because although leaf classification apps have been implemented for leaves in other countries, such an automated app with a reliable identification and classification mechanism has not still been implemented for the purpose of locals. There are existing applications which can identify plants with low prediction accuracies. However, these applications are based on foreign plant data sets that do not include valuable herbs and shrubs with medicinal qualities. Therefore, it is the main research gap which can be identified through the existing researches and implemented automated applications regarding plant identification and classification.

If the main research gap is summarized, it is the "Lack of technological research carried out on the recognition and classification of Ayurvedic plants in Sri Lanka using machine learning and computer vision, and unavailability of a full-automated mobile application for the particular purpose." This is mostly important in the system of medicine called Ayurveda in Sri Lanka, because identification and classification of medicinal plants is considered an important activity in the preparation of herbal medicines. Because treatment using a wrongly identified medicinal plant may claim the life of a patient. Additionally, it is important for Ayurveda practitioners and also traditional botanists to know how to identify and classify the medicinal plants through computers [5]. They should apply their knowledge of Ayurveda with technical approach too.

Following table shows the comparison of several existing mobile applications with plant classification functionality with our system, in the means of functionalities, technologies and techniques used for implementation, devices in which can be used as well as accuracy, which is the most important criteria to be considered.

| Reference | Functionalities of the Application | Technology and techniques Used for implementation | Devices in which can be used | Accuracy |
|---|---|---|---|---|
| **Pl@ntNet [40]** | Image sharing and retrieval for the identification of plants, works with several parts of the plant including flowers, leaves, fruits and bark, allows integrating user's observations in the database to a collaborative workflow involving the members of a social network specialized on plants | Generalist CBIR method for the visual search engine | iPhone and iPad devices | Medium |
| **LeafSnap [41]** | Identification of tree species from photographs of their leaves. | Automatic Visual Recognition, using computer vision components for discarding non-leaf images and for segmentation | iPhone and iPad devices | Medium |
| **MedLeaf [42]** | Medical plant identification based on leaf image and document searching of medicinal plant | Computer vision, machine learning and deep learning, Local Binary Pattern (LBP) to extract leaf texture and Probabilistic Neural Network (PNN) for classification | Android mobile devices | Medium |
| **ApLeafis [43]** | Automatically identifies plant species by the photographs of the tree leaves (digital/captured) photograph | Based on CBIR technique | Android mobile devices | Medium |
| **Arogya – Our System** | Classification of a group of rare ayurvedic plants in Sri Lanka with the use of leaf/digested root/fruit of a particular plant, GIS for plant locations, showcase of recipes and prescriptions from them with accurate details | Computer vision and transfer learning based on deep Convolutional Neural Networks (CNN) | Android mobile devices | Higher accuracy is expected |

Table 1.2.1: Comparison with existing systems and identifying gaps

### 1.3. Research problem

The problem regarding the identification and familiarization of the plants, limited to the specified herbal medicine plant leaves, along with the seen necessity for the development of herbal medicine plant identification should be addressed by a study. Especially in Sri Lanka we have our own set of rare Ayurvedic herbs. But most of us are unable to identify these plants due to lack of knowledge. Though many are aware of the existence of different herbal medicine plants and their familiar applications, many are still unable to identify which are these herbal medicine plants from the vast diversity of plants in the environment, so as their noted and approved applications by health professionals.

There are several methods to recognize a plant. At present, plants are identified manually by taxonomist, which are prone to human errors. But manual identification process requires prior knowledge and also it is a lengthy process. Leaf Identification by mechanical means often leads to wrong identification. Identification and classification of medicinal plants are essential for better treatment. Lack of experts in this field makes proper identification and classification of medicinal plants a tedious task. This is one of the major problems in this domain.

Additionally, researchers in the field of botany, medicine, chemical structure analysis, agriculture, and ayurvedic medicinal practitioners, forest department officials, those who are involved in the preparation of ayurvedic medicines and others who are concerned with plant studies are faced with the application of considerable effort on identifying plants. It is stated that plant identification demands extensive knowledge and uses complex terminology, even professional botanists need to take much time in the field to master plant identification [7]. Identifying a medicinal plant with required medicinal values is one of the major challenging tasks faced by them. It plays a crucial role in the preparation of ayurvedic medicines. But lack of expert taxonomists is a major issue in this area. Even though herbal medicine has no side effects, treatment using a wrongly identified medicinal plant may claim the life of a patient.

Plants are an indispensable part of our ecosystem and the dwindling number of plant varieties is a serious concern. Another major problem is that, there is a growing scientific consensus that plant habitats have been altered and species are disappearing at rates never witnessed before. The biodiversity crisis is not just about the perilous state of plant species but also of the specialists who know them This initially requires data about various plant varieties, so that they could be monitored, protected and can be used for future. Plants form the backbone of Ayurveda and today's modern-day medicine and are a great source of revenue. Due to Deforestation and Pollution, lot of medicinal plant leaves have almost become extinct. So, there is an urgent need for us to identify them and regrow them for the use of future generations. Due to growing illegal trade and malpractices in the crude drug industry on one hand and lack of sufficient experts on the other hand, an automated and reliable identification and classification mechanism in order to handle the bulk of data and to curb the malpractices is needed.

# 2. OBJECTIVES

## 2.1 Main Objective

The primary objective of the proposed solution is to develop an android mobile application that gives user the ability to identify and classify a group of selected important ayurvedic plant species in Sri Lanka, based on photographs of the plant's leaf as well as other parts such as digested root and fruit, taken with mobile phone. And it will be attempted to identify herbal plants using deep learning analysis in order to assist more locals to identify them.

The identification and classification aspect is achieved through an android mobile application. The app uses a photograph of a plant leaf/digested root/fruit captured by the user, or taken from the gallery, to give an output of to which plant species it belongs to. Therefore, anyone without prior knowledge will be able to identify and classify the particular ayurvedic plant hopefully.

The development strategy and methodology used in this approach will be able to be used and extended to identify any ayurvedic herb in Sri Lanka furthermore.

## 2.2 Specific Objectives

1. A newly captured, prepared and annotated Ayurvedic plant dataset from Sri Lanka will be trained on Convolutional Neural Network (CNN) from scratch.

2. The obtained noisy image set of medical leaves/digested roots/fruits will be analyzed using deep neural network architectures (Convolution Neural Network models) based on transfer learning.

    - However, training a model from scratch is too costly. To overcome this challenge transfer learning will be applied on the Ayurvedic dataset.

3. The best Convolution Neural Network architecture, which produces the highest accuracy will be used to produce the final deep learning prediction model, that can be applied effectively for the relevant application.

- According to the above methodology, several CNN architectures will be applied in this study experimentally to obtain the best accurate model, by comparing the accuracies with each.

4. This model will be then used to build a mobile application in the Android platform with TensorFlow-Lite which can identify Ayurvedic plants using the built-in camera module.

    - On a mobile device, the Ayurveda plant detection has to be done with time, battery life critical manner, especially when it has to be done in a forest area. In the proposed system the whole identification process will take place on mobile devices and it doesn’t require internet. Therefore, this will be a great solution to identify Ayurveda plants in deep forest areas, where mobile network coverage is not available.

# 3. METHODOLOGY

## 3.1 Understanding the key pillars of the research domain

The system mainly relies on the key pillars of Deep Learning, Transfer Learning based on deep CNN and Data Augmentation.

### 3.1.1 Deep Learning

According to many previous research works done [10-30], identification and classification of plants have been carried out for the last few decades using classical image processing and classification methods. These methods have used shape, texture and color-based features to perform classification. Aspect ratio, eccentricity, kurtosis, skewness, energy, correlation, sum variance, entropy and compactness are some of these features [20]. Excessive computation time required for handcrafted feature extraction is the major problem associated with these classical methods. Nowadays all the classical methods are replaced by machine learning techniques.

Deep learning is a class of machine learning in which a computational model learns to perform classification tasks directly from images. It is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example [37]. Higher accuracy, ability to handle huge volume of image data, inbuilt ability to use GPUs for parallel computation and availability of inbuilt pre-trained Convolutional Neural Networks contribute towards the popularity of deep learning. Since Arogya performs classification on huge volume of image data, deep learning approach will be the best choice.

### 3.1.2 Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data [39]. A CNN uses a system much like a multilayer perceptron that has been designed for reduced processing requirements. The layers of a CNN consist of an input layer, an output layer and a hidden layer that includes multiple convolutional layers, pooling layers, fully connected layers and normalization layers. The removal of limitations and increase in efficiency for image processing results in a system that is far more effective, simpler to trains limited for image processing and natural language processing [39].

CNN has been recognized as a competent method for image recognition in past few years, on mobile devices in an offline environment. CNNs work somewhat similar to neurons in brain. It also performs image transformation with a certain degree of rotation and distortion [35]. Since this network is avoiding the complex preprocessing of the image, we will able to input the original image of the suspected plant directly.

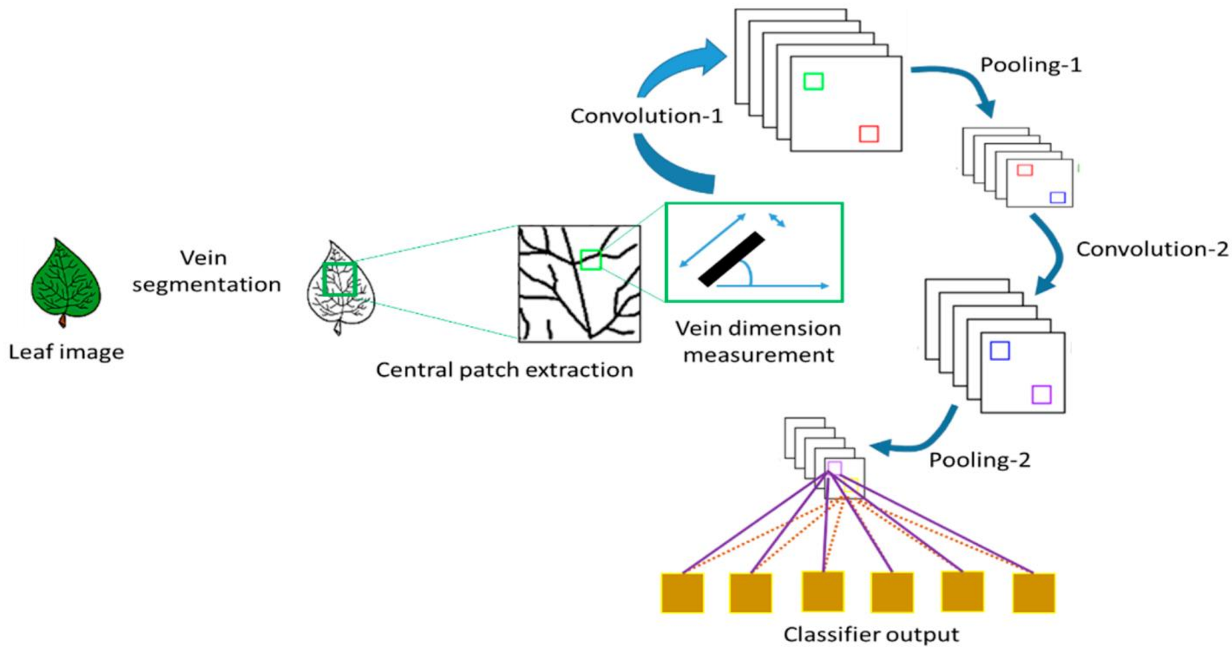


Figure 3.1.2.1: Review on techniques for plant leaves CNN:

Source: https://www.google.com/https://www.mdpi.com

### 3.1.3 Transfer Learning based on deep CNN

An interesting benefit of deep learning neural networks is that they can be reused on related problems. Transfer learning refers to a technique for predictive modeling on a different but somehow similar problem that can then be reused partly or wholly to accelerate the training and improve the performance of a model on the problem of interest [38]. In deep learning, this means reusing the weights in one or more layers from a pre-trained network model in a new model and either keeping the weights fixed, fine tuning them, or adapting the weights entirely when training the model [38].

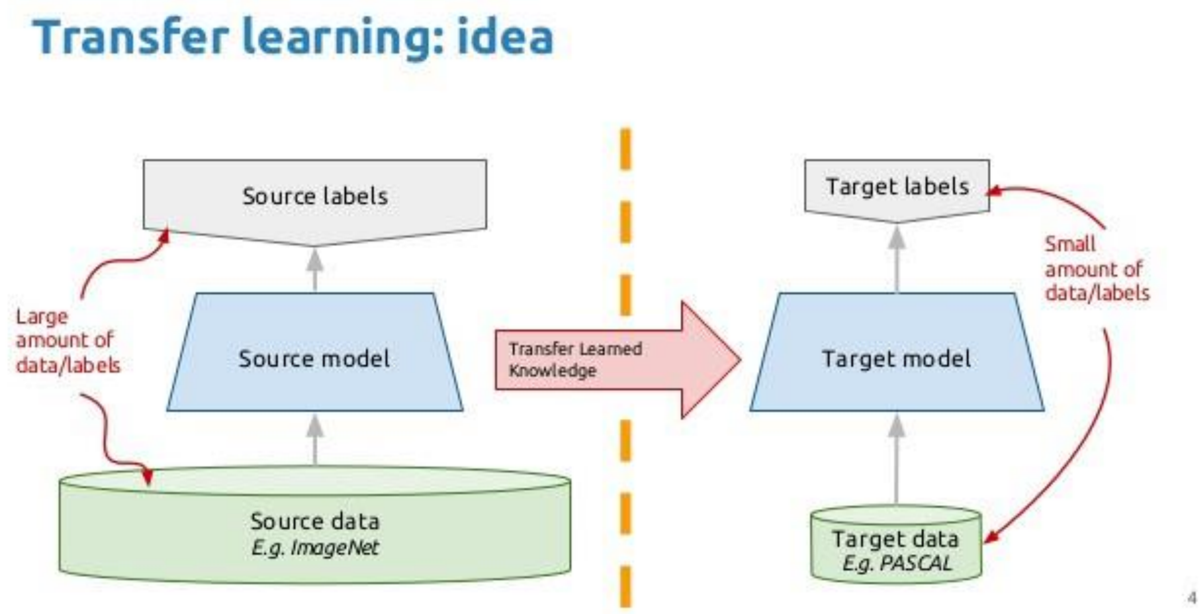


Figure 3.1.3.1: Transfer Learning with Convolutional Neural Networks in PyTorch: Source: https://www.google.com/towardsdatascience.com/transfer-learning-with-convolutional-neural-networks-in-pytorch

Following is the general outline for transfer learning for object recognition [18]:

1. Load in a pre-trained CNN model trained on a large dataset
2. Freeze parameters (weights) in model's lower convolutional layers
3. Add custom classifier with several layers of trainable parameters to model
4. Train classifier layers on training data available for task
5. Fine-tune hyperparameters and unfreeze more layers as needed

### 3.1.4 Data Augmentation

As mentioned in many researches in this domain, among many of DNN structures, the Convolutional Neural Networks (CNN) are currently the main tool used for the image analysis and classification purposes [17]. Although great achievements and perspectives, deep neural networks and accompanying learning algorithms have some relevant challenges to tackle [17]. The most frequently mentioned problem in the field of machine learning, is the lack of sufficient amount of the training data or uneven class balance within the datasets [17]. One of the ways of dealing with this problem is so called data augmentation [17]. For training the convolutional neural network that we use for extracting features, there is an idea for training images to be augmented. As a CNN requires a large number of images to train it, various data augmenting methods is planned to be applied to the collected dataset to increase the size of the dataset.

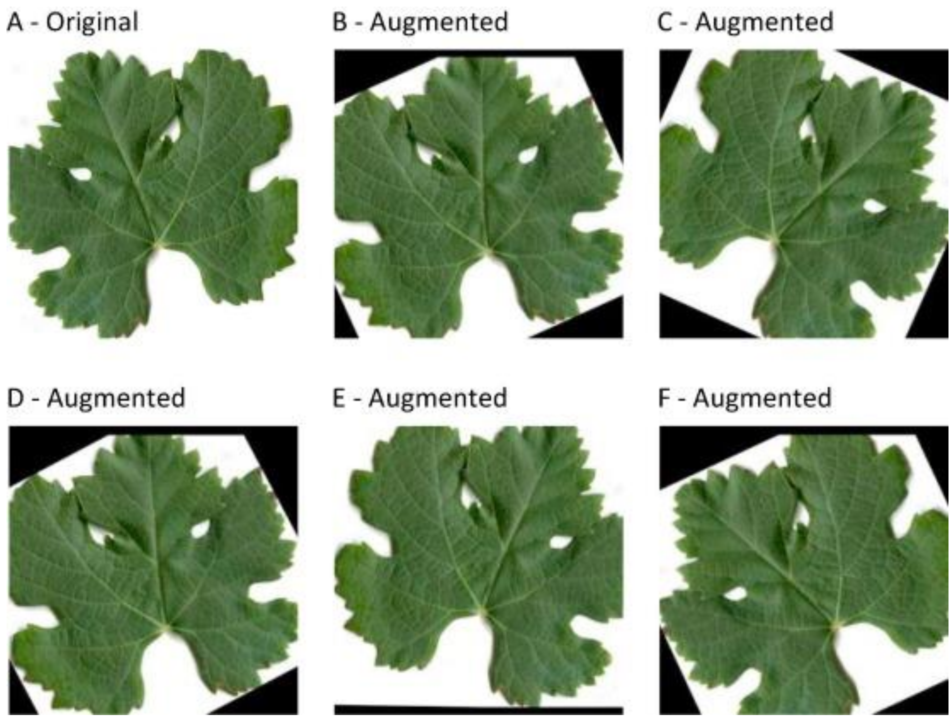


Figure 3.1.4.1: Data Augmentation with leaves:
Source: https://www.google.com/https/www.sciencedirect.com/science/article/pii

### 3.2 Identification and Classification of Ayurvedic Plants in Sri Lanka Using A Mobile Application in An Offline Environment Approach

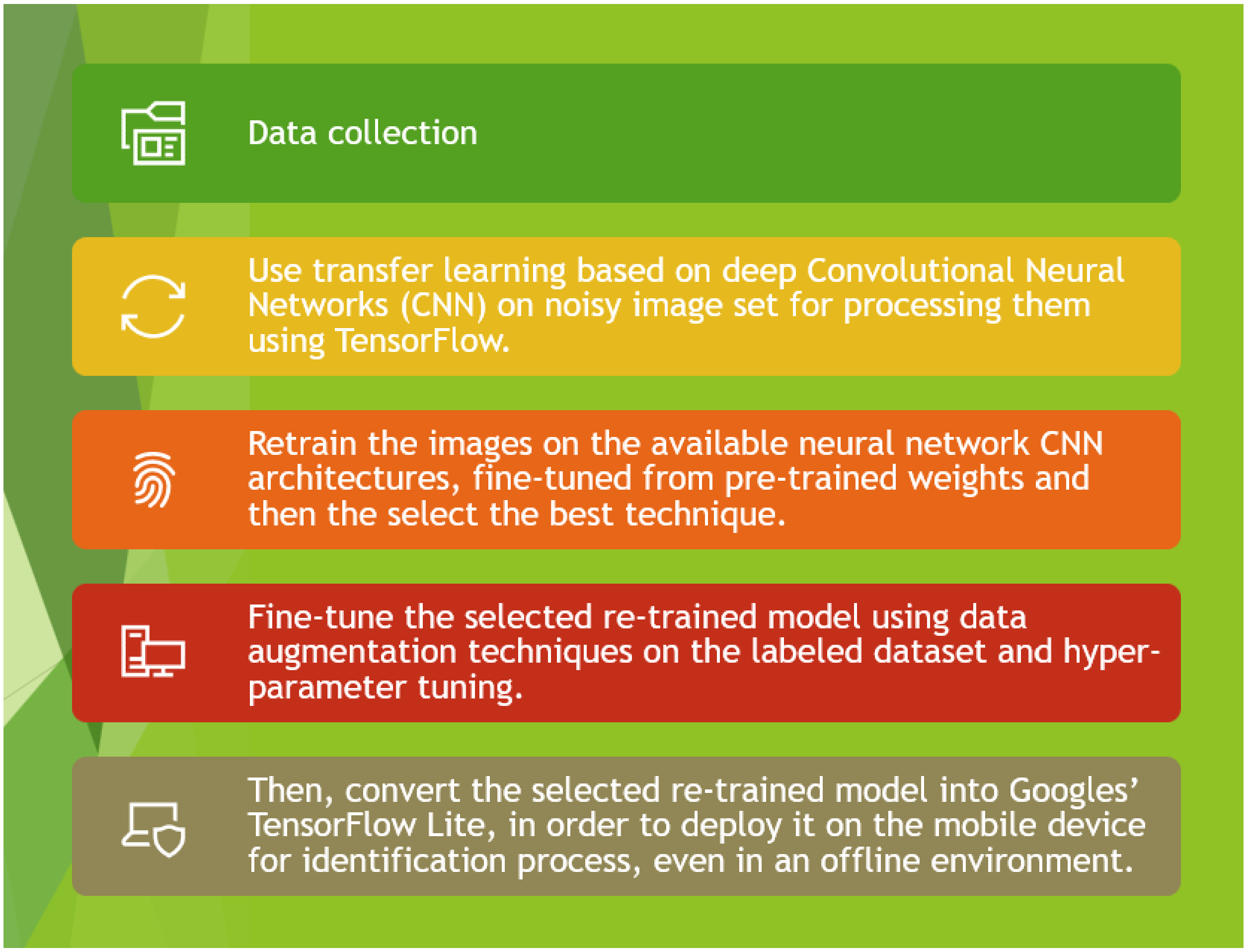


Figure 3.2.1: Methodology of the approach

In this research, the images of the leaf/digested root/fruit that will be scanned using the application will act as the input for this phase. Among many herbal plants, 5 most important ayurvedic plants in Sri Lanka will be chosen to analyze further in detail and the images of the plants will be acquired from plant nursery of Navinna Ayurveda Medical Hospital, social media, Institute of Ayurveda, Alternative medicine website and blogs related to Sri Lankan herbal plants creating a noisy web data set.

After acquiring the images of leaves/fruit/digested root of the selected Ayurveda plants, they will be labelled and annotated using multi-class classification, by forming 5 classes of the particular plants. After the dataset is prepared, transfer learning based on deep Convolutional Neural

Networks (CNN) will be used on noisy image set for processing them using TensorFlow, in a local computer. However, training a model from scratch is too costly. To overcome this challenge transfer learning will be applied in the Ayurvedic dataset. The proposed identification method will be based on running a CNN.

After that, the dataset will be retrained on the available neural network CNN architectures (such as GoogleNet, Inception v2, Inception v4, VGG-16, VGG-19 architectures), then will be fine-tuned from pre-trained weights and then the best technique with the highest accuracy will be selected.

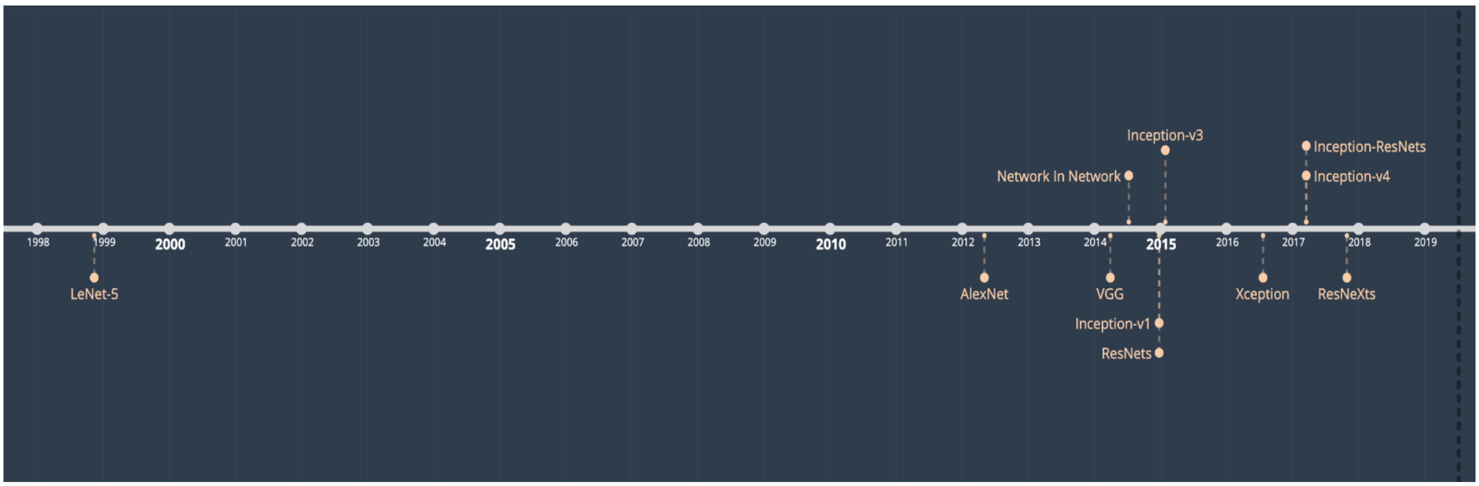


Figure 3.2.2: The 10 architectures and the year their papers were published.
Source: https://towardsdatascience.com/illustrated-10-cnn-architectures-95d78ace614

Then, the selected re-trained model will be fine-tuned using data augmentation techniques on the labeled dataset and hyper-parameter tuning. It will be re-trained with the dataset using Keras which is an open source neural network library, written in python. It is user friendly, modular and extensible since it has been designed for fast experimentations with deep neural networks.

Then the re-trained model will be converted for the deployment on mobile devices with Googles' TensorFlow Lite. TensorFlow Lite is a TensorFlow's lightweight solution for mobile devices which provides on-device machine learning inference with a small binary size and low latency. Tensorflow provides an interface for expressing machine learning algorithms and an application for executing these algorithms [36].

One of the main challenges in detection of Ayurveda plants are similarity between different plant species in different phases of their life cycles and complexity of the background since the end user will be using this application in deep forest environment. Therefore, we cannot just rely on a single feature, such as color, texture or shape to distinguish them. The same species of leaves will be different because of the shades of colors, shape, scale, viewpoint etc. [35].

On a mobile device the Ayurveda plant detection has to be done with time, battery life critical manner, especially when it has to be done in a forest area .In the proposed system the whole identification process will take place on mobile devices and it doesn't require internet .Therefore this will be a great solution to identify invasive plants in deep forest areas, where mobile network coverage is not available .The mobile application will be built using android.

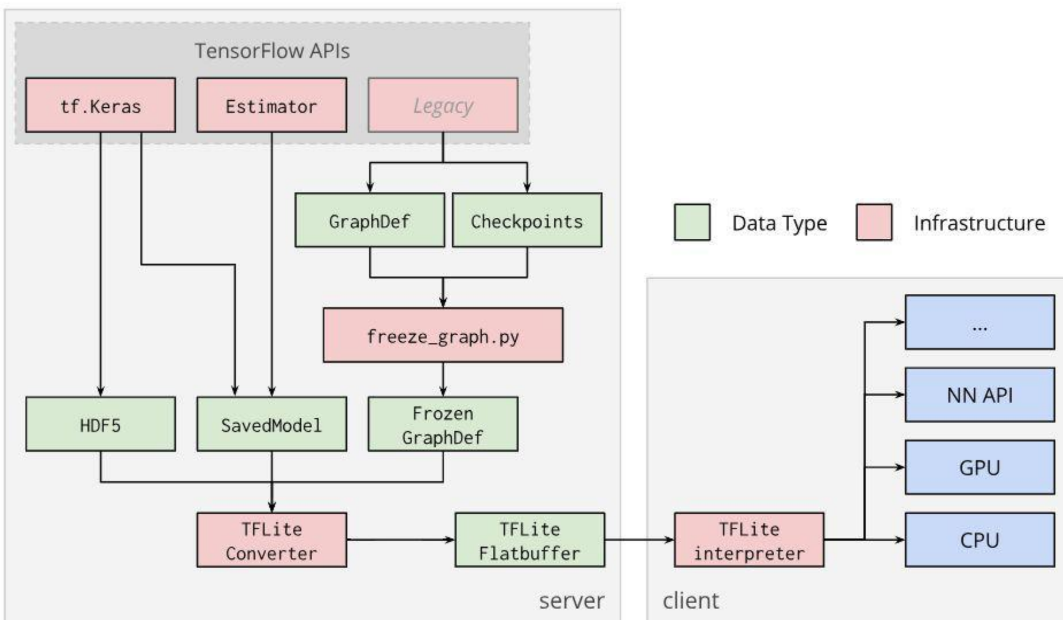


Figure 3.2.3: TensorFlow Lite Converter
Source: [https://www.tensorflow.org/lite/convert](https://www.tensorflow.org/lite/convert)

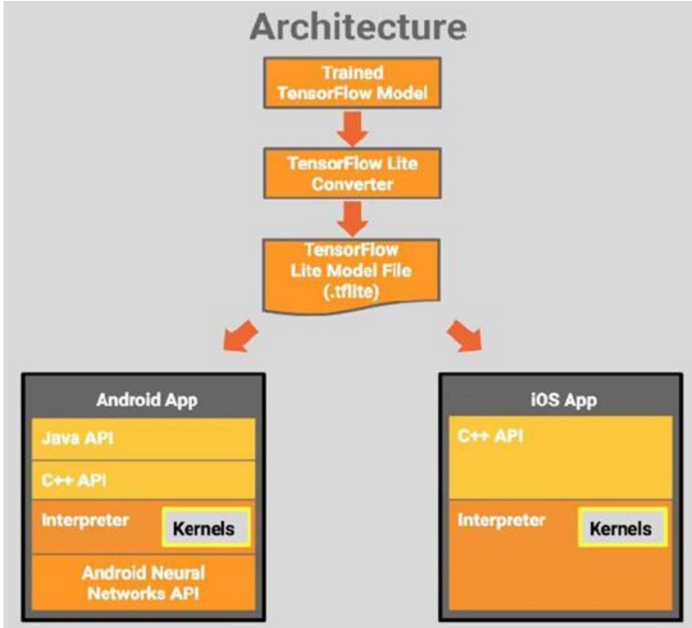


Figure 3.2.4: TensorFlow Lite Architecture
Source: https://www.tensorflow.org/lite/guide

### 3.3 Summary of methodology

The following table summarizes the methodology to be used in the functionality of Ayurvedic leaf classification precisely.

| **Research** | **Research Work** | **Methods and Algorithms to be used** | **Expected Accuracy** | **Remarks** |
|---|---|---|---|---|
| Arogya – An Intelligent Ayurvedic Herb management Platform -><br><br>Classification of a group of rare ayurvedic leaves in Sri Lanka | Use transfer learning based on deep Convolutional Neural Networks (CNN) on collected image set for processing them using TensorFlow<br><br>(The selected re-trained model will be finetuned using data augmentation techniques on the labeled dataset and hyper-parameter tuning.) | CNN Architectures (Such as GoogleNet, Inception v2, Inception v4, VGG-16, VGG-19, etc.) will be used to choose the best technique | Higher Accuracy will be expected | Several existing algorithms will be analyzed, and accuracy will be tested in order to select the suitable algorithms to classify the selected 5 plants accurately. |

Table 3.3.1 : Summary of methodology

## 3.4 High-Level Architecture Diagram

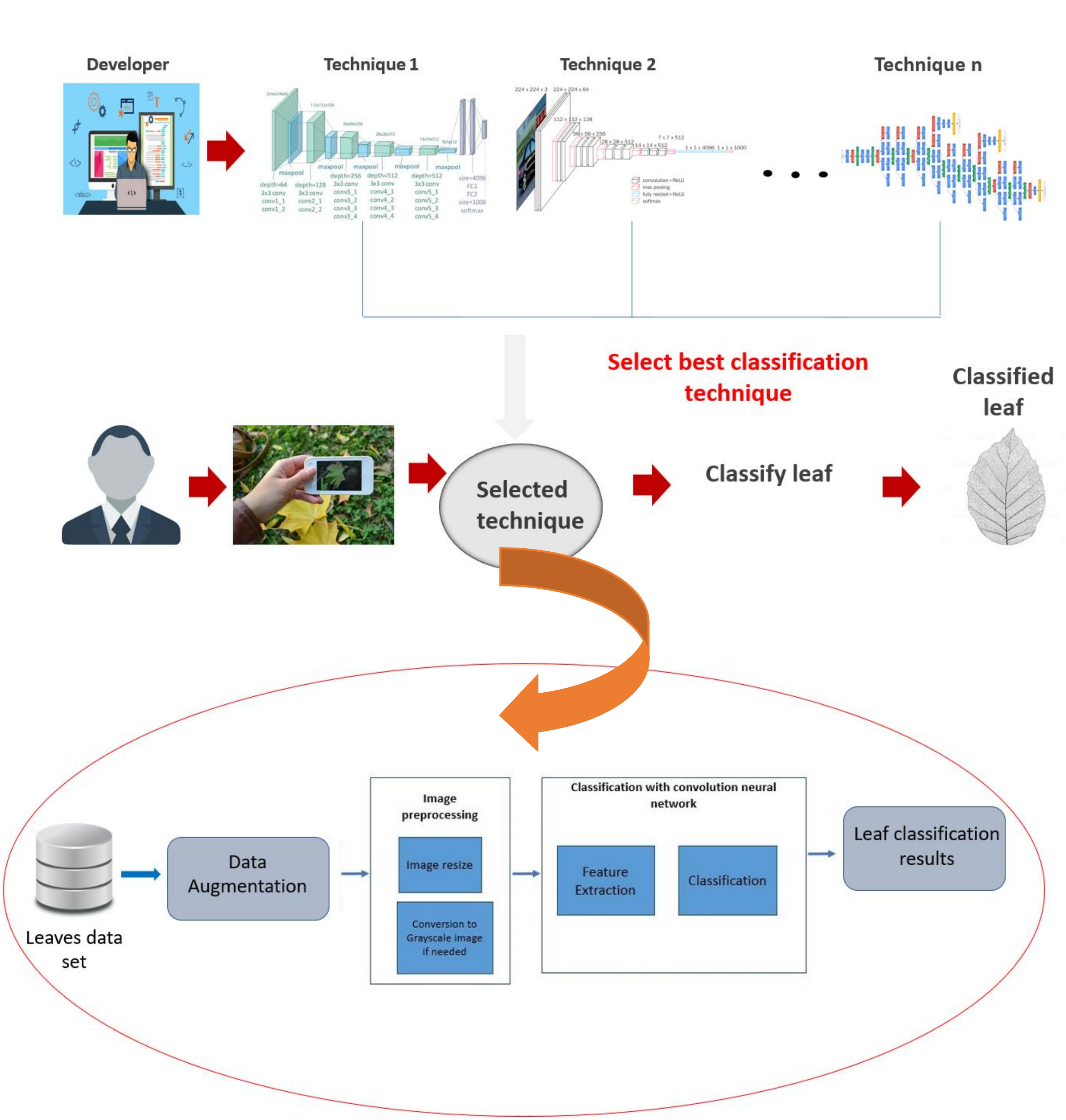


Figure 3.4.1: Software Architectural diagram

# 4. DESCRIPTION OF PERSONAL AND FACILITIES

Facilitators: -

Mrs. Lokesha Weerasinghe - Sri Lanka Institute of Information Technology (SLIIT)

Dr. Dharshana Kasthurirathna - Sri Lanka Institute of Information Technology (SLIIT)

Mr. Sudheera Liyanage - Pearson Lanka (pvt) Limited

Facilities: -

Navinna Ayurveda Medical Hospital

Institute of Ayurveda

Plant nursery of Navinna Ayurveda Medical Hospital

## 4.1 Gantt chart

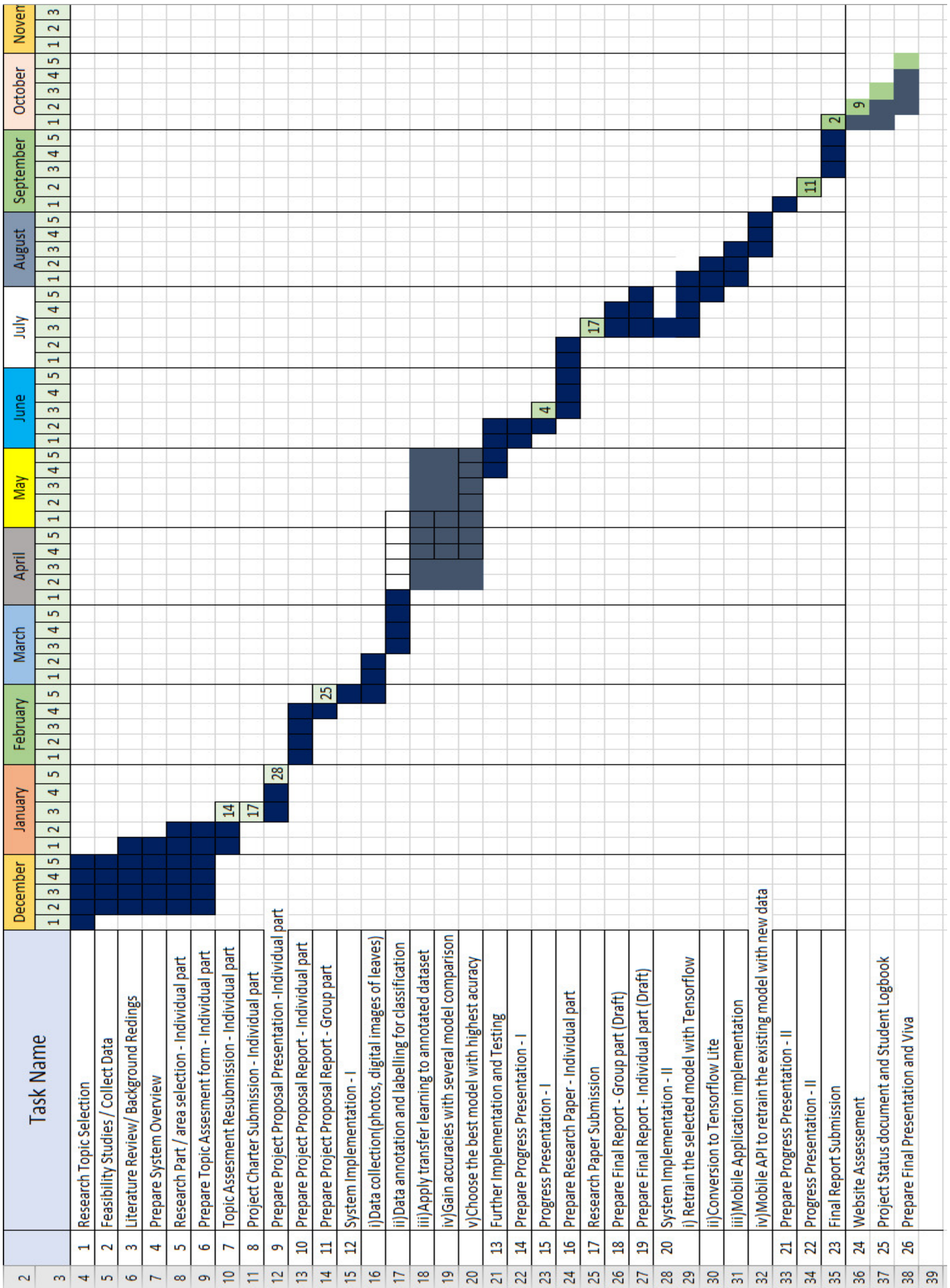


Figure 4.1.1: Gantt chart

## 4.2 Workload Allocation

### 4.2.1 Work Breakdown Structure

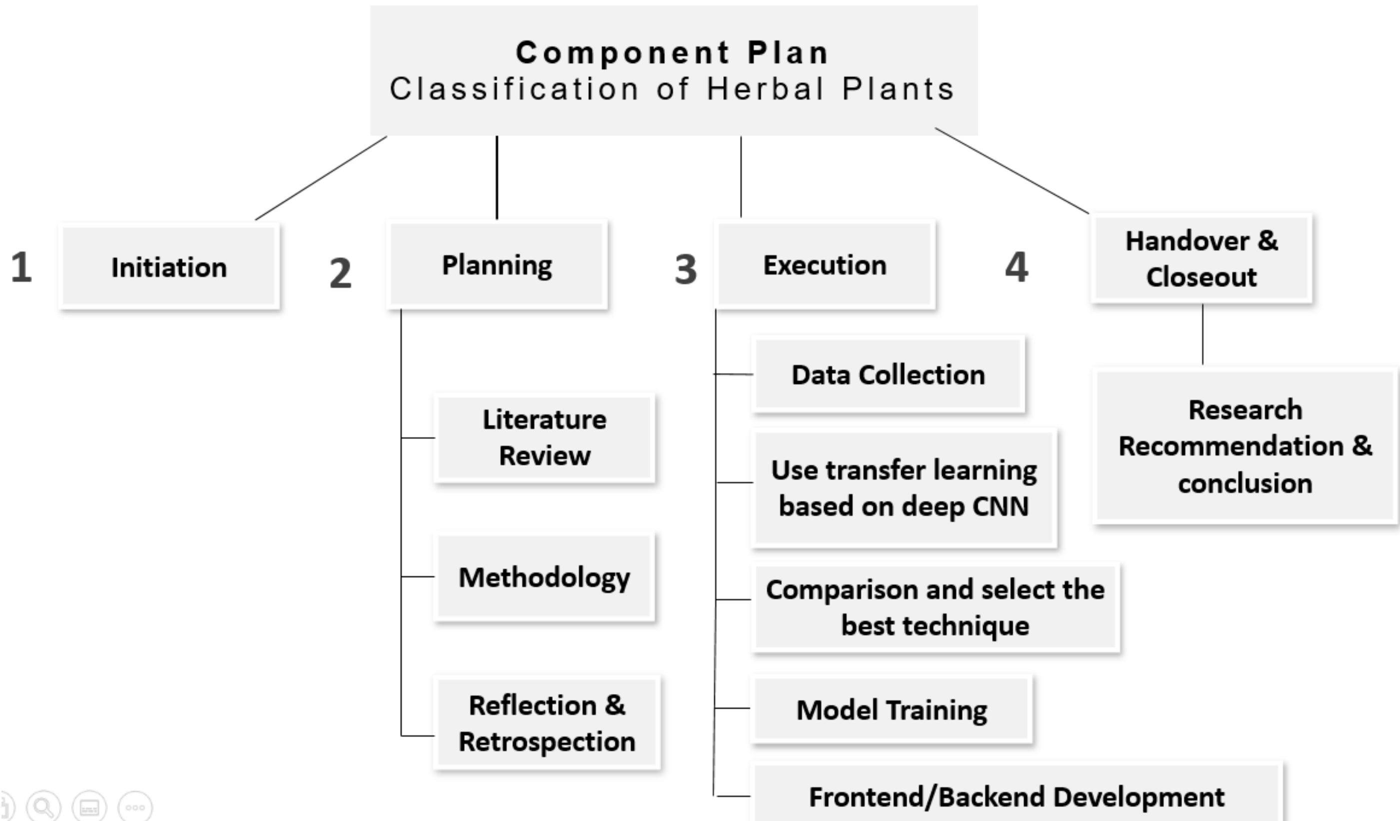


Figure 4.2.1: Work Breakdown Structure

### 4.2.2 Stepwise

1. Collect a good dataset.
2. Use transfer learning based on deep CNN on the image set for processing them with TensorFlow.
3. Retrain the images on the available neural network CNN architectures and then the select the best technique.
4. Re-train the selected model with TensorFlow.
5. Convert the re-trained model to TensorFlow lite.
6. Build mobile application features.
7. Build the mobile API to re-train the existing model with new data.
8. Search functions in mobile client.

## 4.3 Project Requirements

### 4.3.1 Functional Requirements

1. Classify a set of selected Ayurveda herbs in an offline environment.
2. Select the best classification technique with comparing several transfer learning models based on deep CNN.
3. Compare the accuracies of these models with justification.
4. Request to re-train the existing model with new datasets by premium uses.
5. Show classification history on the mobile device.
6. Manage Ayurveda herb details in mobile device.

### 4.3.2 Non-Functional Requirements

1. The application should be able to give the results as fast as possible.
2. Accuracy should be high, and results should be efficient.
3. Feasible in order to understand the functionalities of the app.
4. User friendly Interfaces should be provided.
5. User Experiences should be properly managed in order to achieve a specific functionality.

### 4.3.3 Expected Ways of Testing the Accuracy of the selected predictive models

From all the data collected (captured photos of the tree leaves/digested roots/fruits, digital images, web images, etc.), after the process of data annotation, labelling into 5 classes for the purpose of annotation and totally pre-processing; the prepared dataset will be divided into 3 parts as:

1. 70% of data for training the model – **Training dataset**
2. 15% of data for validating the model – **Validation dataset**
3. 15% of data for testing the model – **Test dataset**

- Training set: A set of examples used for learning that is to fit the parameters of the classifier [45]. In other words, it is used to adjust the parameters of the model, e.g. NN architecture. The aim is to reduce bias or our predictions (i.e. to fit the data) [46].

- Validation set: A set of examples used to tune the parameters of a classifier, for example to choose the number of hidden units in a neural network [45]. In other words, the validation set (also called dev set) is applied for hyper-parameter tuning (regularization, early stopping, drop-off, learning rate...) to reduce variance, i.e. improve the generalization capacity of the particular model (eliminate over-fit). The validation data is not seen by the particular model during training [46].

- Test set: A set of examples used only to assess the performance of a fully specified classifier [45]. In other words, test data is used to provide an unbiased evaluation of a final model. It is not seen by your model at all. Test data should be your real-life data [46].

### 4.3.4 Expected Wireframes for the mobile application to be implemented

Following are some of the wireframes expected to be implemented according to the function of leaf classification in Arogya.

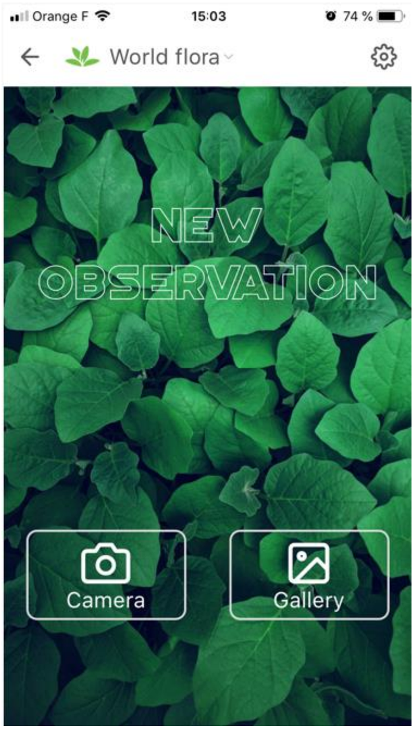


Figure 4.3.4.1:Image capturing interface

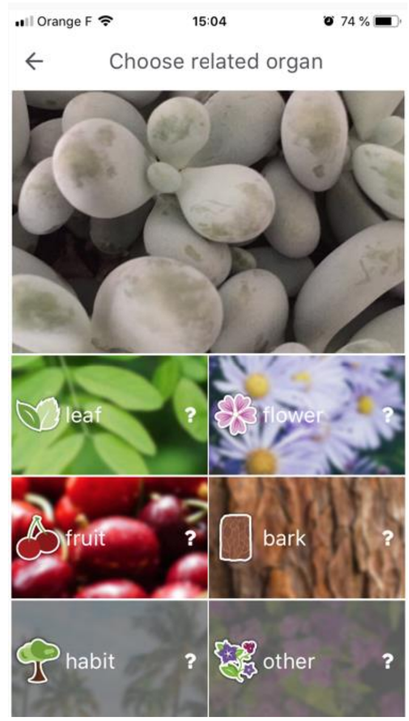


Figure 4.3.4.2: Plant part selection

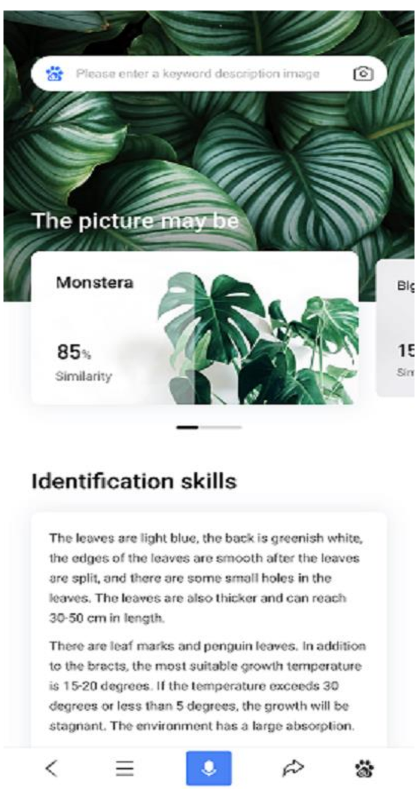


Figure 4.3.4.4:Image classification interface

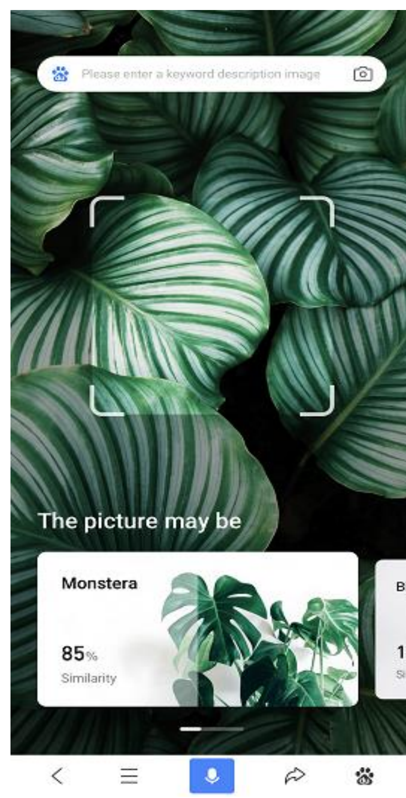


Figure 4.3.4.3:Image focus interface

## 5. TECHNOLOGIES

1. Python
2. Keras
3. TensorFlow
4. TensorFlow Lite
5. Google Colab
6. Flutter
7. Android
8. Firebase – for the database

Research Area – Deep Learning Algorithms (deep neural network architectures based on transfer learning) and Image Processing Techniques

# 6. BUDGET AND BUDGET JUSTIFICATION

| Requirements | Cost |
|---|---|
| **Internet Cost (per month)** | 2000.00 |
| **Documentation printout cost** | 3000.00 |
| **Data collection travelling charges** | 3000.00 |
| **Total Target Cost (per year)** | **30000.00** |

# 7. COMMERCIALIZATION

## 7.1 Target Audience

- People who use ayurvedic treatment
- Researchers in the field of botany, medicine, chemical structure analysis, agriculture, ayurvedic medicinal practitioners, forest department officials, those who are involved in the preparation of ayurvedic medicines and others who are concerned with plant studies
- Doctors, Students, locals and foreigners
- Ayurvedic plant sellers

## 7.2 Market Space

- No age limitations for the users
- No need of advance computer literacy
- No need of advance knowledge in Ayurveda field

## 7.3 Revenue Earning

- Through subscription fee
- Revenue via additional services

# 8.REFERENCES


[1] Pathirage Kamal Perera, "Current scenario of herbal medicine in Sri Lanka", ASSOCHAM, 4th annual Herbal International Summit cum Exhibition on Medicinal & Aromatic Products, Spices and finished products(hi-MAPS), pp. 1-3, April 2012.

[2] Lakpura LLC, Ayurveda. Accessed on: February 15,2020. [Online]
Available: https://lanka.com/about/interests/ayurveda/

[3] Y. R. Azeez, R. A. C. P. Rajapakse, "An application of image processing techniques in identifying herbal plants in Sri Lanka", Proceedings of the 3rd. International Research Symposium on Pure and Applied Sciences, pp.1, 26th October 2018.

[4] Ayurveda Bansko, Blog, Benefits of Ayurveda and the difference to the traditional western medicine, 2015. Accessed on: February 15,2020. [Online]
Available: https://www.ayurvedabansko.com/benefits-ayurveda-and-difference-to-western-traditional-medicine/.

[5] Basavaraj S. Anami, Suvarna S. Nandyal, A. Govardhan, "A Combined Color, Texture and Edge Features Based Approach for Identification and Classification of Indian Medicinal Plants", International Journal of Computer Applications (0975 – 8887), Volume 6– No.12, pp.1-3, September 2010.

[6] Robert G. de Luna, Renann G. Baldovino, Ezekiel A. Cotoco, Anton Louise P. de Ocampo, Ira C. Valenzuela, Alvin B. Culaba, Elmer P. Dadios, "Identification of Philippine Herbal Medicine Plant Leaf Using Artificial Neural Network", IEEE, 978-1-5386-0912-5/17/$31.00,PP.1-2, 2017.

[7] Rademaker C. A. 2000, "The classification of plants in the United States Patent Classification system", World Patent Information 22: 301–307.

[8] Neeraj Kumar, Peter N Belhumeur, Arijit Biswas, David W Jacobs, W John Kress, Ida C Lopez, João VB Soares, "Leafsnap: A computer vision system for automatic plant species identification," ECCV, pp. 502–516. Springer, 2012.

[9] Cem Kalyoncu, Önsen Toygar, "Geometric leaf classification," Computer Vision and Image Understanding", vol. 133, pp. 102–109, 2014.

[10] Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto, Paulus Insap Santosa, "Leaf classification using shape, color, and texture features," arXiv preprint arXiv:1401.4447, 2013.

[11] Thibaut Beghin, James S Cope, Paolo Remagnino, SarahBarman, "Shape and texture based plant leaf classification," Advanced Concepts for Intelligent Vision Systems, 2010, pp. 345–353.

[12] James Charters, Zhiyong Wang, Zheru Chi, Ah Chung Tsoi, David Dagan Feng, "Eagle: A novel descriptor for identifying plant species using leaf lamina vascular features," ICME-Workshop, 2014, pp. 1–6.

[13] Heysem KAYA, Ilhan KEKLIK, Tolga ENSARI, Fatih ALKAN, Yagmur BIRICIK, “Oak Leaf Classification: An Analysis of Features and Classifiers”, 978-1-7281-1013-4/19/$31.00, pp.1-4, 2019.

[14] James S Cope, Paolo Remagnino, Sarah Barman, Paul Wilkin, “The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms,” Advanced Concepts for Intelligent Vision Systems, pp. 135–144, 2010.

[15] Muammer TORKOGLU, Davut HANBAY, “Combination of Deep Features and KNN Algorithm for Classification of Leaf-Based Plant Species”.

[16] Jing Hu, Zhibo Chen, Meng Yang, Rongguo Zhang, Yaji Cui, “A Multiscale Fusion Convolutional Neural Network for Plant Leaf Recognition”, 1070-9908, IEEE SIGNAL PROCESSING LETTERS, VOL. 25, NO. 6, JUNE 2018.

[17] Agnieszka Mikołajczyk, Michał Grochowski, “Data augmentation for improving deep learning in image classification problem”, 17859931, ISBN: 978-1-5386-6143-7, May 2018.

[18] Will Koehrsen, Medium, towards data science, Transfer Learning with Convolutional Neural Networks in PyTorch, November 27, 2018. Accessed on: February 15,2020. [Online] Available: https://towardsdatascience.com/transfer-learning-with-convolutional-neural-networks-in-pytorch-dd09190245ce.

[19] Inside.Techlabs, Leaf Classification Project, October 23, 2018, Accessed on: February 15,2020. [Online]
Available: https://medium.com/@inside.techlabs/leaf-classification-project-89bb5c7577f3.

[20] Dileep M.R, Pournami P.N, “AyurLeaf: A Deep Learning Approach for Classification of Medicinal Plants”, IEEE, 978-1-7281-1895-6/19/$31.00, pp.1-4, 2019.

[21] Jun Woo Lee, Yeo Chan Yoon, "Fine-Grained Plant Identification using wide and deep learning model", Institute for Information & communications Technology Promotion (IITP), No. 2016-0-00010-003, Digital Content In-House R&D.

[22] Benjaphan Sommana, Thanaruk Theeramunkong, “Improving Plant Recognition using Hybrid features from Connectionist and Knowledge-Based Approaches”, IEEE, 978-1-7281-0161-3/18/$31.00, 2018.

[23] Hulya Yalcin, Salar Razavi, “Plant Classification using Convolutional Neural Networks”, Visual Intelligence Laboratory Electronics & Telecommunication Engineering Department Istanbul Technical University, 2017.

[24] Jyotismita Chaki, Ranjan Parekh, Samar Bhattacharya, “Recognition of Whole and Deformed Plant Leaves using Statistical Shape Features and Neuro-Fuzzy Classifier”, IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS), 978-1-4799-8349-0/15/$31.00, 2015.

[25] Chaoyun Zhang, Pan Zhou, Chenghua Li, Lijun Liu, “A Convolutional Neural Network for Leaves Recognition Using Data Augmentation”, IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable,

Autonomic and Secure Computing; Pervasive Intelligence and Computing, DOI 10.1109/CIT/IUCC/DASC/PICOM.2015.318, 978-1-5090-0154-5/15 $31.00, 2015.

[26] Sue Han Lee, Paul Wilkin, Chee Seng Chan, Paolo Remagnino, "DEEP-PLANT: PLANT IDENTIFICATION WITH CONVOLUTIONAL NEURAL NETWORKS", IEEE, 978-1-4799-8339-1/15/$31.00, 2015.

[27] ArunPriya C, Balasaravanan T, Antony Selvadoss Thanamani, "An Efficient Leaf Recognition Algorithm for Plant Classification Using Support Vector Machine", Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering, 978-1-4673-1039-0/12/$31.00, 2012, March 21-23, 2012.

[28] Shanwen Zhang, YouQian Feng, "Plant Leaf Classification Using Plant Leaves Based on Rough Set", 2010 International Conforence on Computer Application and System Modeling (ICCASM 2010), 978-1-4244-7237-6/10/$26.00, 2010.

[29] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang, Qiao-Liang Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", IEEE International Symposium on Signal Processing and Information Technology, 978-1 -4244-1 835-0/07/$25.00 , 2007.

[30] Pushpa BR, Anand C, Mithun Nambiar P, "Ayurvedic Plant Species Recognition using Statistical Parameters on Leaf Images", International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 11, Number 7 (2016) pp 5142-5147.

[31] Sethulekshmi A V, Sreekumar K, "Ayurvedic leaf recognition for Plant Classification", Sethulekshmi A V et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6), ISSN:0975-9646, 2014, 8061-8066.

[32] A Gopal, S. Prudhveeswar Reddy, V. Gayatri, "Classification of Selected Medicinal Plants Leaf Using Image Processing", IEEE, 978-1-4673-2322-2112/$31.00, 2012.

[33] J. W. Tan, S. Chang, S. Binti Abdul Kareem, H. J. Yap, K. Yong, "Deep learning for plant species classification using leaf vein morphometric", pages 1–1, 2018.

[34] R. Janani, A. Gopal, "Identification of selected medicinal plant leaves using image features and ann", 2013 International Conference on Advanced Electronic Systems (ICAES), pages 238–242, Sept 2013.

[35] Gavai, N. R., Jakhade, Y. A., Tribhuvan, S. A., & Bhattad, R. (2017, December). MobileNets for flower classification using TensorFlow. In 2017 International Conference on Big Data, IoT and Data Science (BID) (pp. 154-158). IEEE.

[36] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Kudlur, M. (2016). Tensorflow: A system for large-scale machine learning. In 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16) (pp. 265-283).

[37] MathWorks, Deep Learning. Accessed on: February 19,2020. [Online]
Available: https://www.mathworks.com/discovery/deep-learning.html.

[38] Machine Learning Mastery, How to Improve Performance With Transfer Learning for Deep Learning Neural Networks. Accessed on: February 19,2020. [Online]
Available: https://machinelearningmastery.com/how-to-improve-performance-with-transfer-learning-for-deep-learning-neural-networks/.

[39] TechTarget, Convolutional Neural Network. Accessed on: February 19, 2020. [Online]
Available: https://searchenterpriseai.techtarget.com/definition/convolutional-neural-network

[40] Hervé Goëau, Pierre Bonnet, Alexis Joly, Julien Barbe, Souheil Selmi, Vera Baki´c, Jennifer Carré, Daniel Barthelemy, Nozha Boujemaa, "Pl@ntNet Mobile App", Proceedings of the 21st ACM international conference on Multimedia, DOI: 10.1145/2502081.2502251, ppt. 1-2, October 2013.

[41] Neeraj kumar Kumar, Peter N. Belhumeur, Arijit Biswas, João V. B. Soares, "Leafsnap: A Computer Vision System for Automatic Plant Species Identification", European Conference on Computer Vision, DOI:10.1007/978-3-642-33709-3_36, ppt. 1, October 2012.

[42] Desta Sandya, Yeni Herdiyeni, "MedLeaf: Mobile Application For Medicinal Plant Identification Based on Leaf Image", International Journal on Advanced Science, Engineering and Information Technology, Vol. 3 (2013) No. 2 ISSN: 2088-5334, ppt 1-3, May 2013.

[43] Lin-Hai Ma, Zhong-Qiu Zhao, Jing Wang, "ApLeafis: An Android-Based Plant Leaf Identification System", International Conference on Intelligent Computing, LNCS 7995, pp. 106–111, 2013.

[44] PredictiveAnalyticsWorld.com, 3 Ways to Test the Accuracy of Your Predictive Models. Accessed on: February 22, 2020. [Online]
Available:http://https//www.predictiveanalyticsworld.com/machinelearningtimes/3-ways-test-accuracy-%20predictive-models/3295/

[45] Machine Learning Mastery, What is the Difference between Test and Validation Datasets? Accessed on: February 22, 2020. [Online]
Available: https://machinelearningmastery.com/difference-test-validation-datasets/

[46] ResearchGate, What is training and testing data in machine learning?
Accessed on: February 22, 2020. [Online]
Available:https://www.researchgate.net/post/What_is_training_and_testing_data_in_machine_learning