4. Proposed System Design

The experiment is to classify images into 26 classes from Indian Medicinal Leaves Dataset using different transfer learning models. For the purpose of experimentation, dataset were augmented. The pre-trained models namely VGG-16, VGG-19, MobileNet V1 and ResNet 50 have been trained, using the state-of-art systems, on the ImageNet dataset which consists of 1000 classes of images. Instead of using the hand-crafted feature extractors, we have made use of these pre-trained models for the same purpose. Accuracy is used as an evaluation metric shown in Eq 1.

$$Accuracy = \frac{Correct\ test\ set\ predictions}{Total\ No\ of\ Test\ set\ predictions} \quad (1)$$

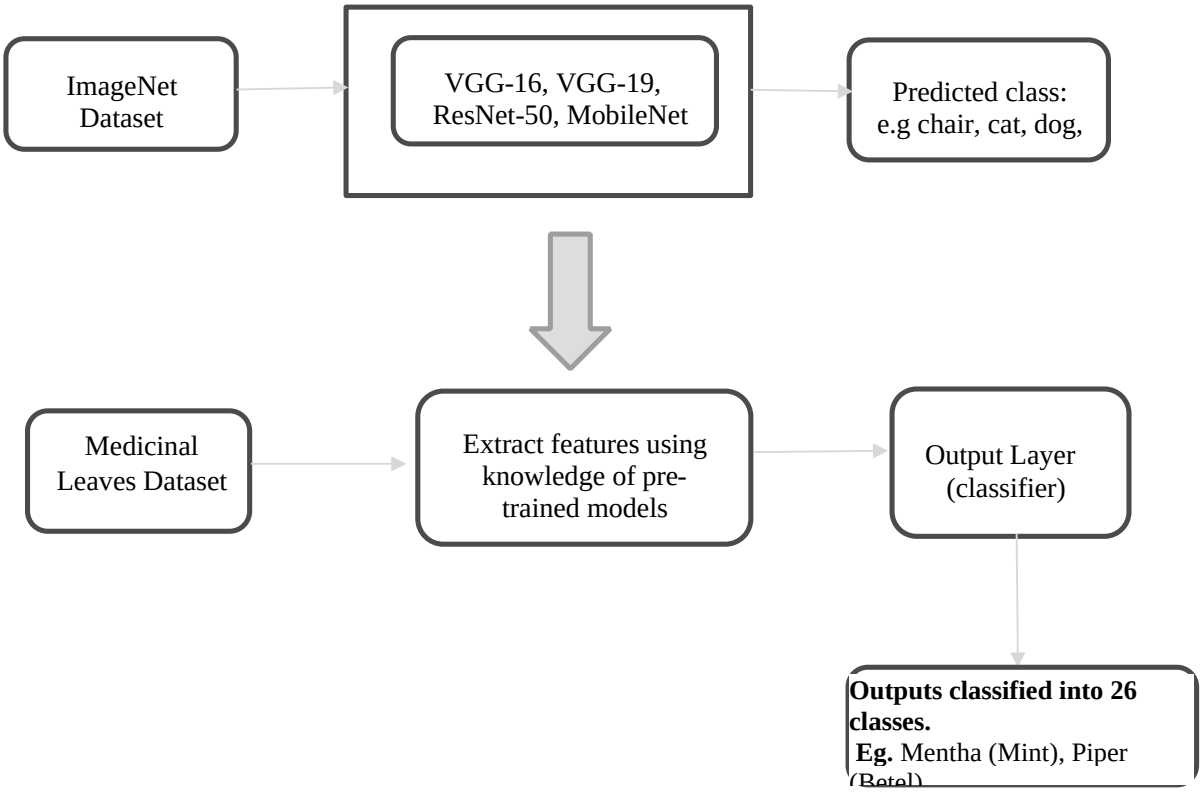

Figure 2: Proposed system overview

4.1. Datasets:

The datasets used in this paper are Indian Medicinal leaves dataset consisting of 26 classes . Only whole images were used. Training set images from all datasets were augmented by applying shear range of 20% and zoom range 20%, translating the images horizontally. As it has been seen in the previous works that CNNs learn better with a greater number of examples and the number of images in some classes of leaves were fairly less. One sample image from each dataset is shown in Figure.

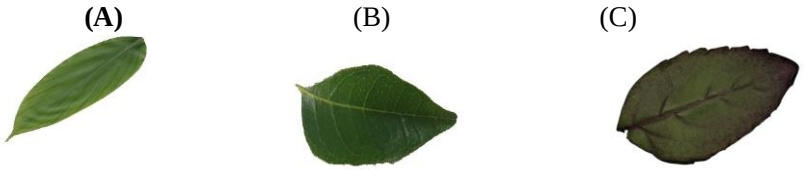

Figure 3 (A) Alpinia Galanga (Rasna) (B) Murraya Koenigii (Curry) (C) Ocimum Tenuiflorum (Tulsi)

4.2. Transfer learning for the feature extraction

Feature set obtained from leaf images can be based on shape, texture, color or venation. These features can also be used in combination. Here, we present a comparison of feature extraction capabilities of 4 pre-trained CNN models. Transfer learning is the concept of using the knowledge acquired by the network using the state-of-art process, to be applied for identification of other related images. In this work it is shown that the features obtained from these pre-trained networks are also efficient in recognizing other related images. Feature extraction helps to curb over-fitting of CNN models, speed up training, improve accuracy and improve data visualization. It also reduces the features from the dataset by creating new features from the existing ones and discarding the original ones. This process creates a summary of the original features from the dataset. Features for our experiments were obtained from fully connected layer of the pre-trained models. They are summarized in Table 1(Depth refers to the topological depth which includes activation layers, batch normalization layers etc).

Table 1: List of pre-trained models used, the input image size, depth and number of features extracted from each model.

Model name	**Input image size**	**Depth**	Trainable parameters
VGG-16	224x224	16	14,714,688
ResNet-50	224x224	50	23,534,592
VGG-19	224x224	19	20,024,384
MobileNet-V1	224x224	28	3,206,976

4.4. Experiments and Results

Various experiments were conducted using different pre-trained models on the mentioned datasets. The ones which helped to achieve exemplary results are shown in this section. The images were first augmented and then resized according to requirements of the feature extractor model used. The size of feature vector for different models is shown in Table 1. The feature vector was then flattened and passed into to the output layer to classify image into one the 26 classes. For the purpose of evaluation, accuracy/ precision was used as a metric. The explanation of the datasets and comparative analysis is provided below.

4.4.1 Evaluation on dataset. Medicinal leaves dataset was introduced It consists of 26 classes of leaves. An image from this dataset is shown in Figure 3 (a). Images of training set are augmented so that the model could learn as many numbers of features as possible. The images are first resized and pre-processed based on the requirements of the feature selector and then classified. Training was done for 50 epochs, loss used here was categorical cross entropy, optimizer used was 'adam' and accuracy as metric and their accuracies obtained on the test set are shown in the Table 2 below.:

Table 2: Accuracy attained from each pre-trained model on Swedish Dataset.

Feature selector	**Accuracy (%)**
VGG-16	97
ResNet-50	68
MobileNetV1	98
VGG-19	96