Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

# Artificial Intelligence: For the Recognition of Plant Species in the Malvaceae Family

**Aruna S.I.[1], Jyothi C.J.[2], Sonu Suresh[3], Sreethumol[4], Sheethal M.S.[5], Ashna Prince[6], Albert Babu[7]**

(1) Assistant Professor (Computational Biology), Department of Botany & Computational Biology, St. Aloysius College, Elthuruth.

(2) Assistant Professor (Botany), Department of Botany & Computational Biology, St. Aloysius College, Elthuruth.

(3,4,5,6,7) Department of Botany & Computational Biology, St. Aloysius College, Elthuruth.

***(Received: 27 September 2025 Revised: 05 October 2025 Accepted: 14 October 2025)***

**KEYWORDS**

Malvaceae, NumPy, scikit-learn, Machine Learning

**ABSTRACT:**
This study aimed to classify leaves of selected plants from the Malvaceae family using convolutional neural networks and deep learning techniques. We utilized 170 images from three species within this family for training. By exploiting CNNs with deep learning, we enhanced efficiency and classification accuracy, significantly reducing training time by using pre-trained models. The study was conducted with Python and various libraries, including NumPy, scikit-learn, and OpenCV. Our CNN model successfully distinguished between the different leaf images, achieving an accuracy of 71%. This approach highlights the benefits of transfer learning, particularly with limited data, and emphasizes the importance of leaf characteristics—such as shape, texture, and venation—in accurate classification. The results showcase the advancements deep learning has made in visual recognition tasks, suggesting future work could expand the dataset, explore other advanced architectures, and apply data augmentation techniques to improve accuracy. Ultimately, this research contributes valuable insights for ecological studies, agriculture, and conservation efforts by aiding in the identification and monitoring of plant species.

## INTRODUCTION

In biology, categorizing plants and organisms is vital because it allows species to be organized according to their similarities and differences. Due to the potential of misidentification, traditional plant classification relied on human competence, which presented hazards in therapeutic contexts. Computer-based plant classification has been made possible by the development of digital technology (Cerutti et al., 2013), which allows for more accurate identification and lowers errors. These days, machine learning and artificial neural networks (ANNs), which imitate the brain's pattern recognition, are frequently employed for these tasks (Salee et al., 2019). In order to properly identify inputs, like pictures of plants, ANNs learn by modifying internal connections during training (Sheikh et al., 2023).

This study focuses on classifying plants in the Malvaceae family, which includes species like cotton, okra, and hibiscus, by assessing leaf traits using deep learning and artificial neural network (ANN) technology. Members of this widely distributed, morphologically diverse family have both medicinal and economic importance. Plants in the Malvaceae family are used in traditional medicine and offer fiber and oils.

The family of flowering plants known as the Malvaceae, or mallows, is thought to have 244 genera and 4225 species (Zhenghao and Meihua 2017). The family's most well-known members include cocoa, cotton, okra, and hibiscus. This family of plants consists of shrubs or herbs that typically have stellate hairs. This family is widely distributed and can be found in a variety of environments worldwide, including desert plants and tropical rainforests (Das et al., 2021). The genus Malva, which contains a number of species referred to as mallows, is whence the word "Malvaceae" originates. The morphology of Malvaceae species varies greatly, encompassing shrubs, trees, and herbs (Das et al., 2021). Flowers occur in a variety of colors and

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

designs, and leaves can be palmately lobed or plain.. Sida acuta, Hibiscus rosa-sinensis, and Abelmoschus esculentus. Members of this family include Gossypium herbaceum and Theobroma cacao. According to Xuet al. (2017), they typically display a tap root system. The plants in this family have woody or herbaceous stems that are branching, upright or spreading, and pubescent with stellate hairs. ERARSLAN & KOÇYİĞİT (2019) describe the leaves as alternating, petiolate, stipulate, stipules deciduous simple, whole or palmately divided or lobed, border wavy or serrate, apex acute, and multicostate reticulate venation. The corolla's characteristics include being enormous, showy, twisted, polypetalous, and occasionally slightly connate at the base with the staminal tube. There are noticeable veins on the petals. The stamens in the Malvaceae family are monadelphous, indeterminate, and form a staminal tube.

Some genera have highly tomentose seed coats, however most seeds are non-endospermic. Common characteristics of plants in the Malvaceae family include the presence of mucilage in plant parts, alternating, simple, stipulate leaves, the presence of epicalyx, and the absence of corolla (Tsopgniet al., 2019). The Malvaceae family is extremely important economically. The stem of Hibiscus cannabinus yields a highly valued fiber that is used to make floor matting, rope, cordage, coarse canvas, and sacs. For this reason, it is used in place of jute. Cotton seed is also significant because it yields a semi-drying edible oil that can be used to make soap and oleomargarine. Additionally, the oil cake makes a very nutritious cattle feed.

Additionally, the residue is used to make nitroglycerine, oil cloth, synthetic leather, and washing powder. The stem of Hibiscus sabdariffa yields a robust, smooth brown fiber that works well as a jute alternative in the paper and textile industries. Fibers of lesser significance are also produced by other hibiscus species. Majangafibres are produced by H. tiliaceous. Okra fibers are produced by Abelmoschus esculentus. Abutilon theophrasti produces Indian mallow or China jute. China cultivates it widely. It is used to make rugs and absorbs dyes rapidly. In all tropical nations, Urena lobate is a weed that produces a yellowish fiber that is often used in place of jute since it lasts longer.

It is used to make coffee bags in Barazil. In addition, they have some therapeutic qualities. It is possible to substitute plasma with this mucilage. Syphilis is treated with an infusion of the roots. In Nepal, cuts, wounds, and boils are treated externally using the root juice (Das & Islam 2019). Althaea officinalis roots have a strong demulcent effect and are extremely rich in mucilage. According to Habtemariam (2019), hibiscus flower buds are cooling and tringent, which helps to relieve body burning, urine discharges, seminal weakness, and piles.

The main goal of this work was to create an artificial intelligence (AI) tool that uses deep learning techniques to correctly identify and categorize a subset of plants from the Malvaceae family. These plants' captioned field photos were used to train the tool. The tool improves automated plant identification by analyzing, recognizing, and classifying plant photos when they are entered. This is done using the knowledge gathered from the trained and tagged field photographs.

The AI tool that was created as a result of this study shown proficiency in identifying Hibiscus rosa sinensis, Sida acuta, and Abelmoschus esculentus, among other members of the Malvaceae family. Since no previous research has examined this particular problem, this categorization approach, which emphasizes morphological traits, closes a present research gap. Future researchers in this subject are expected to greatly benefit from the study's findings, which will serve as a useful resource for additional research.

## MATERIALS AND METHODOLOGY

### PYTHON

Python is a widely-used and widely-supported language with a vast community and numerous powerful tools for scientific computing (De Smedt &Daelemans 2012).. Packages like NumPy, Pandas, and SciPy are freely available and well-documented, making iteration fast. Python is forgiving and allows programs to appear as pseudo code, making it useful for testing pseudo code. However, it is dynamically written and has issues like Duck writing, which can lead to errors and errors in the documentation. (Lutz 2001)

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

**NUMPY**

Numpy is python package which provide scientific and higher level mathematical abstractions wrapped in python. It is the core library for scientific computing, that contains a strong ndimensional array object, provide tools for integrating C, C++ etc. It is additionally useful in linear algebra, random number capability etc. Numpy's array type augments the Python language with an efficient data structure used for numerical work. Numpy additionally provides basic numerical routines, like tools for locating Eigenvector (Harris *et al*., 2020).

**SCIKIT LEARN**

Scikit-learn could be a free machine learning library for Python. It features numerous classification, regression and clustering algorithms like support vector machine, random forests, and k-neighbors', and it additionally supports Python numerical and scientific libraries like NumPy and SciPy (Hao *et al.*, 2019). Scikit-learn is especially written in Python, with some core algorithms written in Python to get performance. Support vector machines are enforced by a python wrapper around LIBSVM .i.e., logistic regression and linear support vector machines by a similar wrapper around LIBLINEAR (Pedregosa *et al*., 2011).

**THE STEPS OF THE PROPOSED SYSTEM**

**Data Collection and Dataset Preparation**

**Data Collection**

Images of the Malvaceae family had been collected from the Thrissur location. The vast majority of the photos of hibiscus, okra, and sida were taken from several Thrissur locations. This can help the model learn under various imaging settings and attributes.

**Dataset Preparation**

In order to train and test the model, data preparation was required. A training set of manually classed and labeled photos was created from the acquired image data, and a testing set of randomly chosen, unclassified, and unlabeled images was created. The photographs in the training set were not the same as the images in the testing set. 119 (70%) of the 170 total photos that were gathered were used as training samples.

**Image Acquisition**

The camera was used to take pictures of the plant leaf. In order to develop a color transformation structure for the RGB leaf image, this image was in RGB (Red, Green, and Blue). A device-independent color space transformation was then implemented.

**Image processing**

Various filtering techniques are used to apply to the input image in the preprocessing phase in order to improve the image quality. The first stage in any image processing application that can help obtain a more meaningful interpretation of a picture is the preprocessing task. Furthermore, preparation operations might impact a classifier's overall performance by exposing obscure or ambiguous information in a picture. Numerous subtasks make up this task, such as image type conversion for computational ease, image enhancement for visual ease, noise removal for consistency in the process, image scaling and resizing or shrinking for quick computation, the use of filter functions, and other crucial methods required to improve the quality of the input leaf image.

**Image Pre-processing**

Various techniques for pre-processing were taken into consideration in order to eradicate noise from the image. Image scaling was used for computational convenience because huge image files can occasionally degrade a model's overall performance. The information contained in the original image could not be abused by the manner this scaling task was applied. Another subtask under the pre-processing task is histogram equalization, which is used to enhance contrast and extract detailed information from the input image. The input image's color space was initially RGB/BGR, but it can be changed to other color spaces for other uses. Many image processing methods perform well with grayscale conversion because grayscale images are easier to process than RGB images.

Most algorithms handle input images differently because different color spaces have varying numbers of potential pixel values. Furthermore, compared to RGB photos, HSV/HIS images made it easier for the algorithm to determine which color was dominant. Here, the picture dilation approach—also referred to as a postprocessing technique at times—was used to restore pixels that had been lost as a result of earlier techniques, primarily segmentation. In order to obtain the intended unbiased information, dilation aids in filling in these missing pixels; nevertheless, it may also introduce extraneous pixels. Occasionally, these compounds are discovered to be sounds by accident. Therefore, denoising techniques such as median filtering are used to eliminate these disturbances.

**RGB to Gray Converter**

According to the weighted technique, also known as the luminosity method, red has the longest wavelength of the three colors, while green has the shortest wavelength and also has a calming impact on the eyes. It implies that we must reduce the red color's contribution, raise the green color's contribution, and place the blue color contribution in between. New grayscale image = (0.3 * R) + (0.59 * G) + (0.11 * B) is the new equation that results. This equation shows that Blue has contributed 11%, Green has provided 59%, which is more than the contributions of the other three colors, and Red has contributed 30%.

**Image Resize**

Document images usually have a resolution of more than 2000 x 2000, which is too big to feed into a CNN given the computer power available today. In addition to requiring more computational resources, a large input dimension increases the likelihood of overfitting. Once an RGB image has been converted to grayscale, it is resized into a standard format, which is 400 Ã 400 for higher resolution. The technique of blurring an image using a Gaussian function—named for the mathematician and physicist Carl Friedrich Gauss—in image processing produces a Gaussian blur, sometimes referred to as Gaussian smoothing. In graphics software, it is a commonly used effect that usually lowers visual noise and detail.

**Image segmentation based on Otsu's method**

In the majority of computer vision applications, image segmentation is a crucial step. Image segmentation is the process of breaking an image up into several areas or segments according to certain standards. Specifically, Otsu's method is among the most widely used techniques for image segmentation. One common method used for image thresholding is Otsu's method. Based on the grayscale intensity values of its pixels, it divides an image into two classes: foreground and background. Additionally, Otsu's method finds the ideal threshold value that divides two regions with the highest inter-class variance by using the grayscale histogram of a picture.

**Feature extraction**

Features may seem as a straightforward numerical function of the picture. In other words, a feature is any technique that extracts a number from the image. These should ideally have purpose. The features were divided into two categories: local and global. The entire image determines the global features. On the other hand, local features are dependent on a local image region and have a position. The uniformity of colors and patterns in an image is referred to as texture. The adjacency idea in images is used by the Gray Level Co-occurrence matrix (GLCM).

The primary concept is that it continuously records the values of neighboring pixel pairs that occur throughout an image. The orientation and separation of image pixels can be used to build the co-occurrence matrix. Certain gray levels that occur periodically control texture patterns. As a result, a particular texture can be identified by repeating the same gray levels at predetermined relative places.

**Machine Learning**

The goal of machine learning is to create algorithms that let computers learn from data and generate precise predictions. Training and prediction are the two primary tasks involved. Through a process called inductive or concept learning, the system uses a dataset (training data) to build a model by finding patterns in the data. In general, a better model is produced with more training data. The model maps fresh inputs to intended outputs during the prediction phase. Ensuring high generalization—that is, making accurate predictions on unseen data—is the primary issue.

**Clustering Algorithm**

Data points are grouped using clustering when they are more similar to one another than to those in other groupings. By gradually merging smaller, related clusters into bigger ones, hierarchical clustering creates a nested structure. Non-hierarchical clustering, on the other hand, uses heuristics to identify effective—though

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

not always ideal—groupings by dividing data into non-overlapping groups without a hierarchical framework. In general, non-hierarchical approaches are more computationally efficient than those that are hierarchical.

CNN was employed in the current study to classify data. Learning a hierarchy of feature detectors and training a nonlinear classifier to recognize intricate document layouts was the major goal. Down sampling and pixel value normalization should be done first on a document image before feeding the normalized image into a CNN to get the class label.

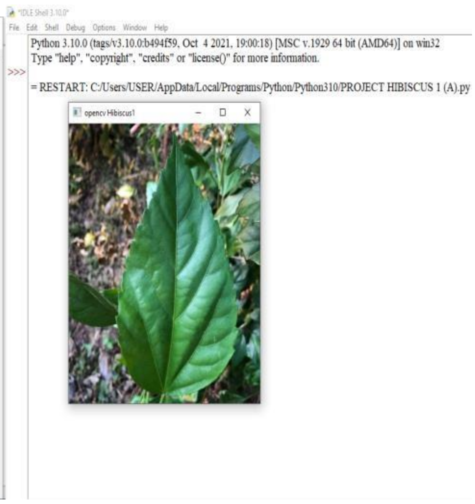
**RESULT**

Image processing was the main part of the work, the images in the prepared dataset will be processed under different techniques for further walk of the implementation. Image Acquisition was the very first step that feeds an input image (original leaf image) to the model, these images can be acquired from the prepared data sets that are collected leaf images from the agricultural field. Initially a dataset was created. The dataset containing the images of plants in Malvaceae family. Mainly the dataset contains the images of *Abelmoschus esculentus, Hibiscus rosa-sinensis,* and S*ida acuta.* The dataset created for the current study was available in figure 1. The created datasets were loaded to python using open cv module for the further processing of the images and was represented in figure: 2. The preprocessing stage plays an important role in many image processing applications here in this research work also preprocessing is used to get better result, this stage can include number of sub stages like image enhancement, noise removal with mean and adaptive median filtering, image scaling and resizing, histogram equalization using filter functions and other important techniques that were necessary to improve the quality of the original leaf image. Segmentation stage was also another essential part of a system in which it helps to focus only on the desired region out of the original leaf image, in the acquired image there may exist some unnecessary parts that are captured with the camera, so the segmentation process was done to get those portions of the preprocessed image which have certain value of pixels. There exist several types of image segmentation techniques in the area of image processing and computer vision having their own advantages and disadvantages according to the problem they are applied on, i.e., the image to be segmented. Thresholding, Fuzzy C Means, Watershed, K means Clustering, and other many, can be listed as an example. Morphological process was an image post-processing technique done based on shape of the object to be processed. It was normally performed on binary images. Therefore, the binarisation of the images were done and was represented in figure :3 after segmentation based on region detection was performed there may exist unnecessary parts still. These unwanted parts can be occurred due to different reasons during the time of image capturing. In the implementation grayscale dilation, erosion or combination of them i.e., opening or closing techniques can be used based on the image under process and was included in figure :4. The feature extraction and selection from an image plays a critical role in the performance of any following classifier. After segmentation and morphological processing, set of features are going to be extracted from the resultant image. In feature extraction stage every image was assigned a feature vector to identify it. This vector is used to distinguish the image. Higher accuracy of the classifier can be achieved by the selection of good feature extraction techniques. Since an image was a large data set use of all the pixel values in classification can create computational overhead as a result features with distinguishing properties should be extracted and then be selected. The details regarding the feature extraction were mentioned in figure: 4 to figure 10. Classification was the major decision-making stage of image recognition. It uses the features extracted in the previous stage to identify the image segment according to preset rules. There are several techniques for classification, some of them are called, Linear Regression (LR), Logistic Regression (LogR), k-Nearest Neighbor (k-NN), Naïve Bayes (NB), Decision Trees (DT), Artificial Neural Networks (ANN), Linear Classifier (LC), Support Vector Machine (SVM), Random Forest (RF) and so forth. In this study, the decision tree was used for the classification purpose. The entire image processing and classification was carried out with the aid of high level, programming language python and various modules and libraries available in python. The training accuracy was represented in figure: 11.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

**Figure:1Data set of Malvaceae family**


**a.*Sidsacuta***

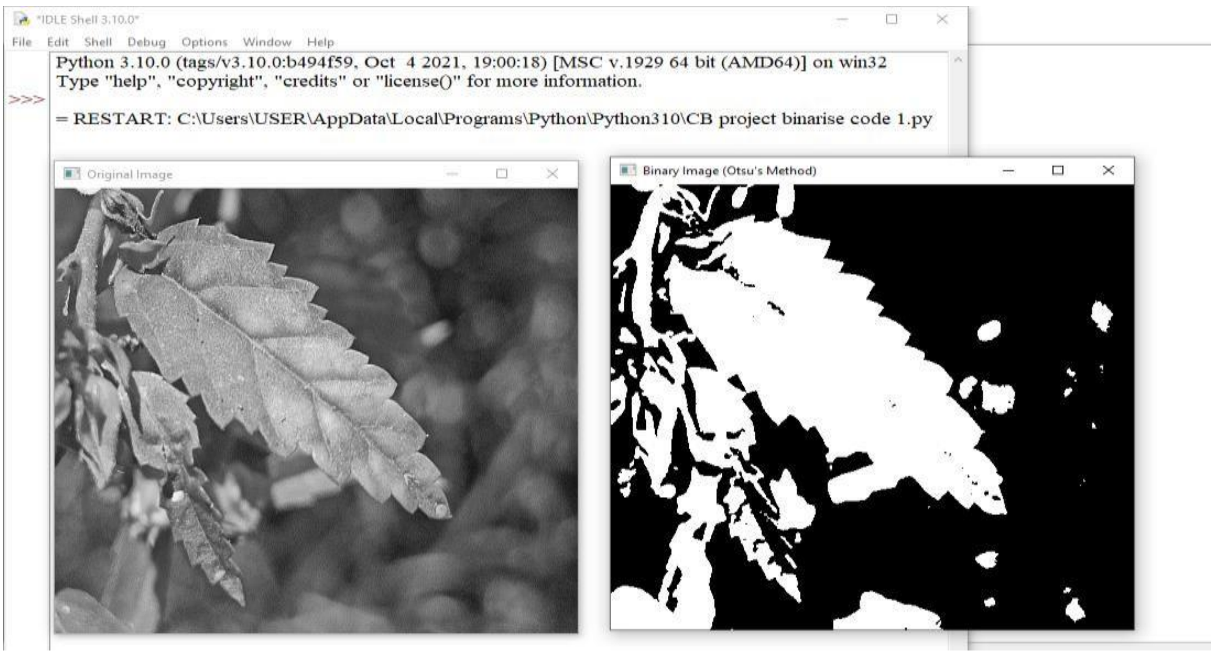



**b.*Hibiscus rosa-sinensis***

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

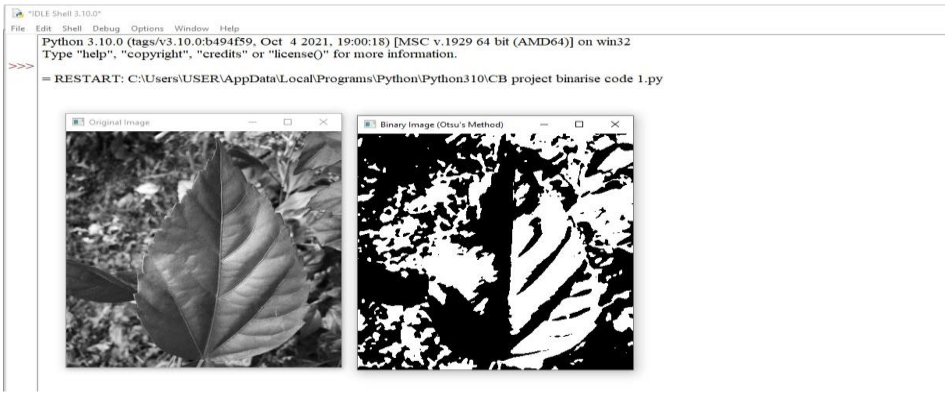

*c. Abelmoschus esculentus*

**Figure: 2 Data sets loaded to python using opencv for image processing. The selected plantspecieswasmentionedasa.*Sidaacuta*,b.*Hibiscusrosa-sinensisandc.Abelmoschusesculentus***

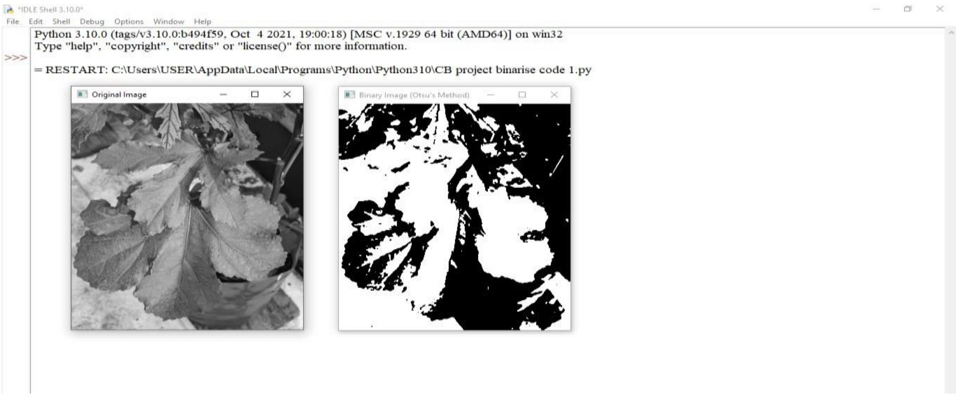



**a. Sidaacuta**

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727



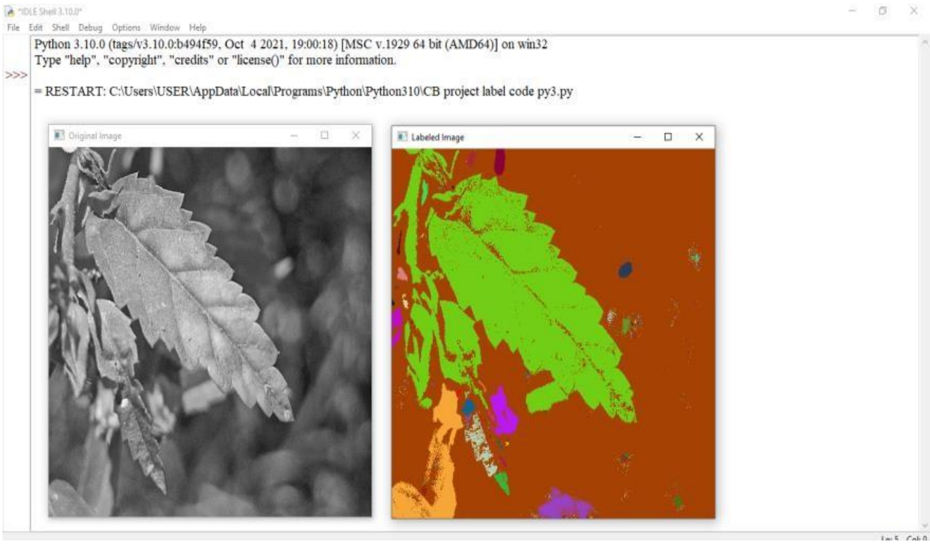

***Hibiscusrosa-sinensis***




***Abelmoschusesculentus***

**Figure3BinarizationofthethreeselectedplantspeciesinMalvaceaefamily.**

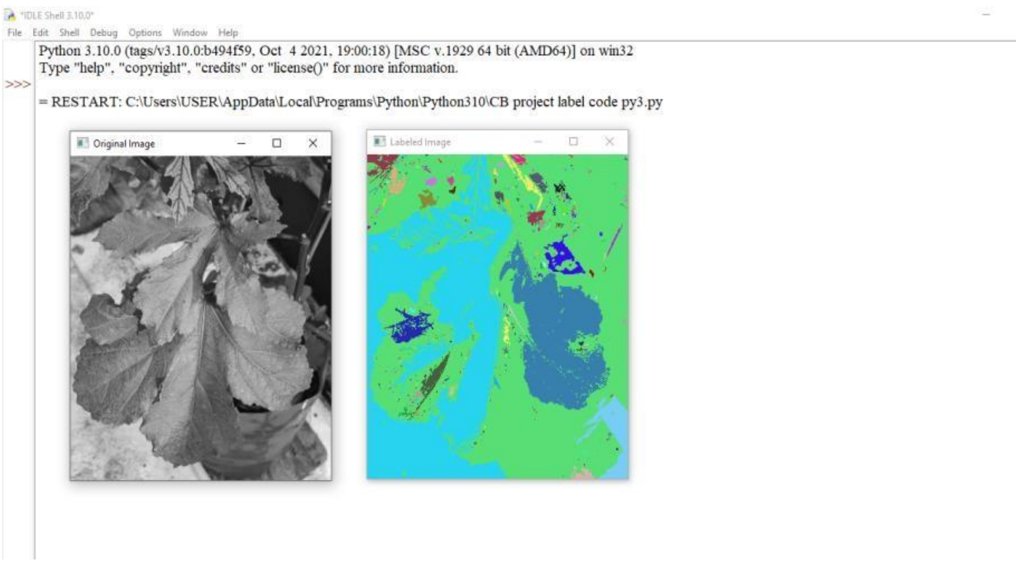
**a.*Sidaacuta*,b.*Hibiscusrosa-sinensisandc.Abelmoschus esculentus***




**a.** ***Sidaacuta***

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727


***b. Hibiscusrosa-sinensis***




***c. Abelmoschusesculentus***

**Figure4:DifferentcolourconversionsofselectedplantspeciesinMalvaceaefamily.**

**a.*Sidaacuta*,b.*Hibiscusrosa-sinensisandc.Abelmoschus esculentus***

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

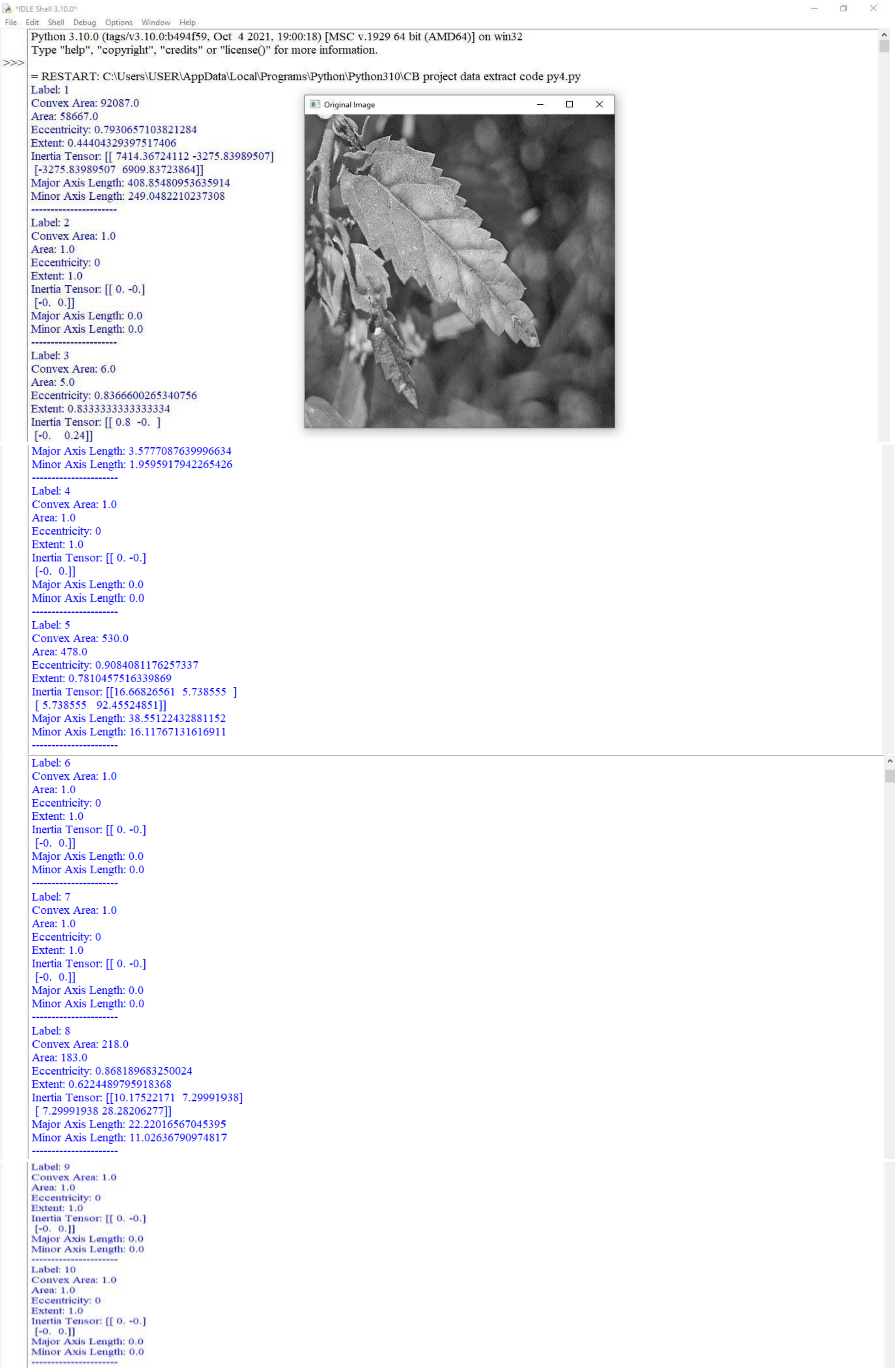



**Figure5Featureextractionof*Sida acuta***

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

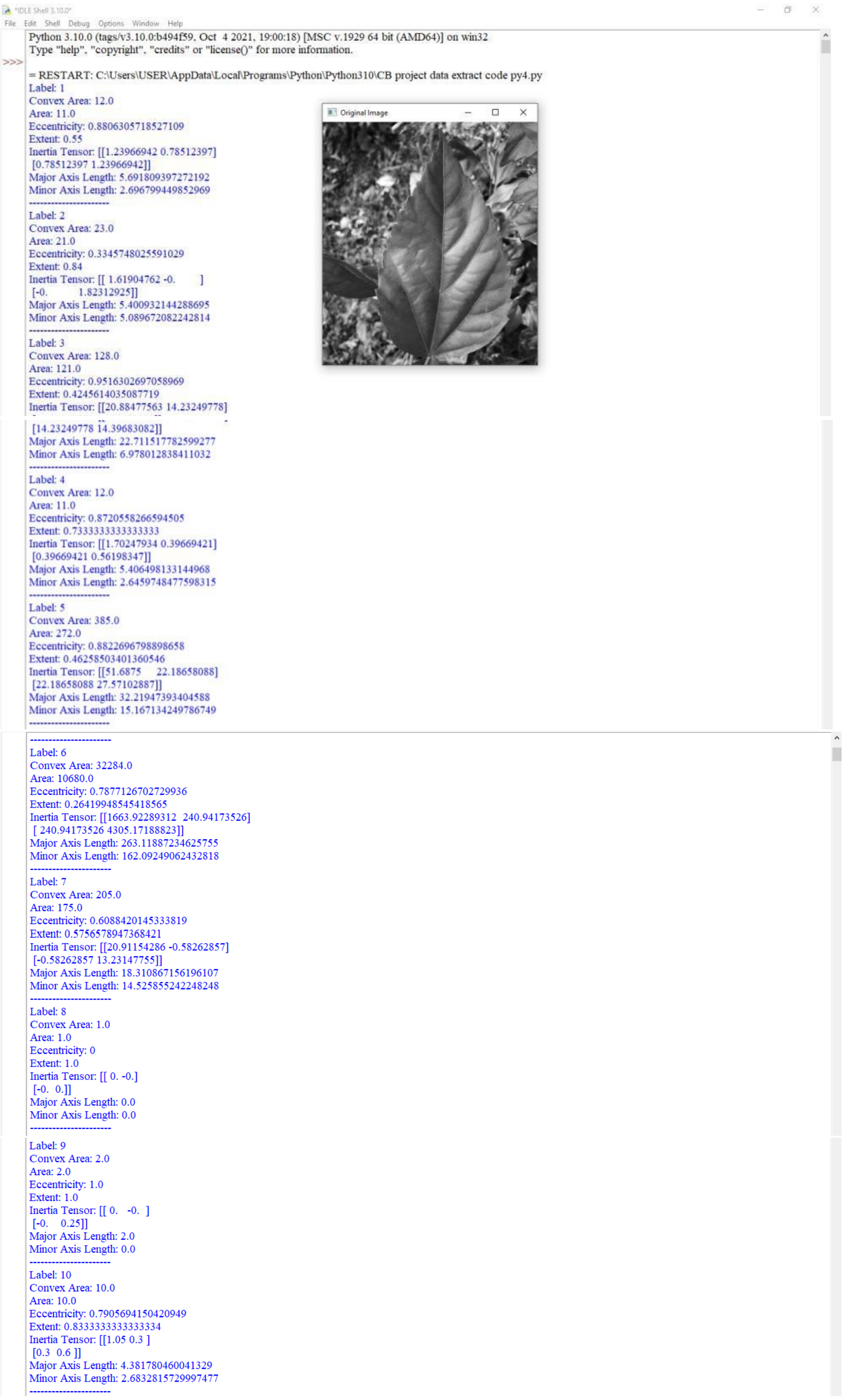



**Figure6Featureextractionof*Hibiscusrosa-sinensis***

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727


**Figure7Featureextractionof*Abelmoschusesculentus***

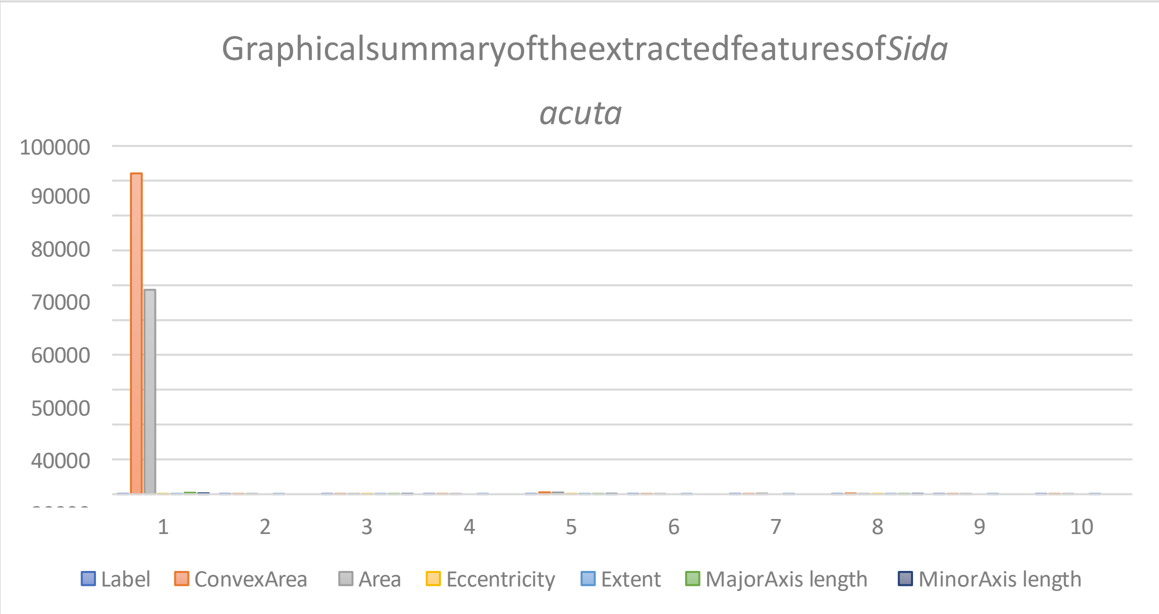



**Figure:8Graphicalsummaryofextractedfeaturesof*Sida acuta***

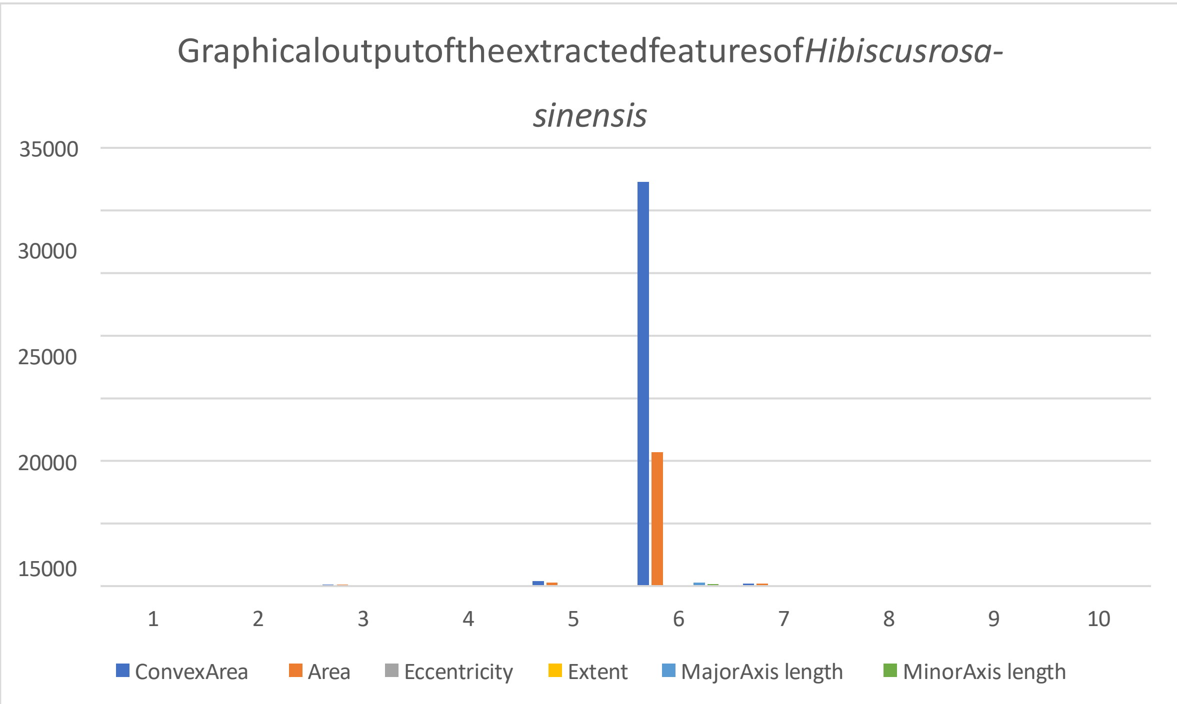



**Figure:9Graphicalsummaryofextractedfeaturesof*Hibiscusrosasinensis***

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(5), 3274-3290 | ISSN:2251-6727

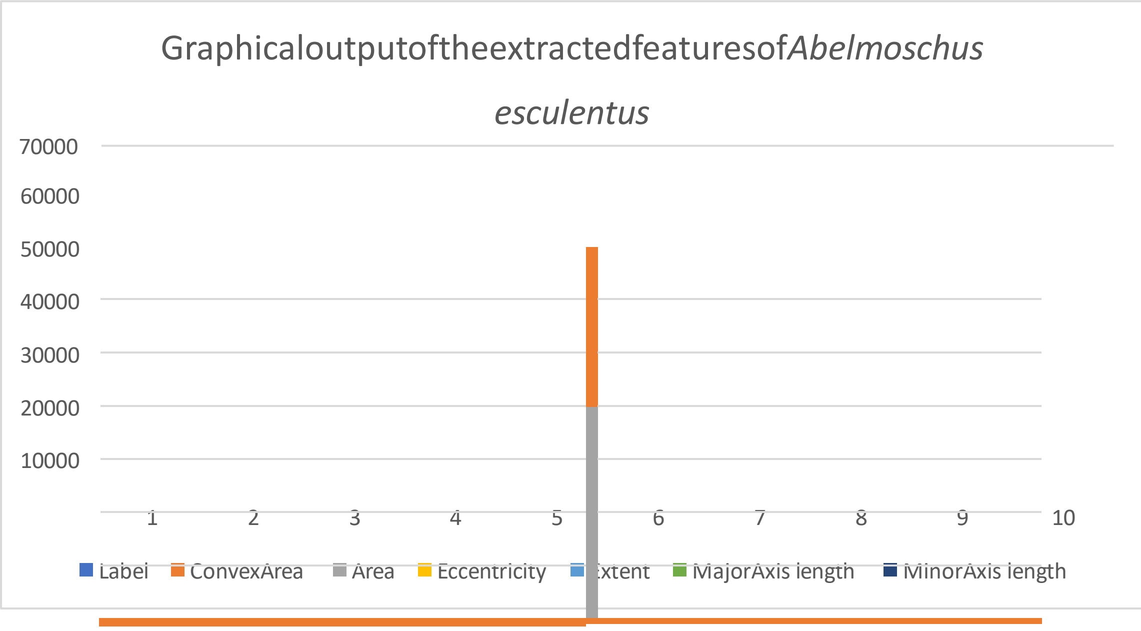



**Figure:10Graphicalsummaryofextractedfeaturesof*Abelmoschusesculentus***


**Figure 11: Accuracy obtained from the trained datasets**

**DISCUSSION**

Image processing has become an invaluable tool in plant detection and classification offering nondestructive and efficient methods for recognizing plant. By analyzing digital images of plant researchers and farmers can easily identify the taxonomic details and features of plants, and make conclusive decisions regarding taxonomy. The process involves acquiring digital images, preprocessing, segmenting, extracting relevant features, and employing classification algorithms. Too many researches carried out to identify the plant species using convolutional neural networks. In an early work, Arafat *et al*. (2016) presented a comparison of three well-known techniques for leaf classification. The authors compared the abilities of these techniques to differentiate among plant species. The evaluated techniques included the Histogram of Oriented Gradient (HOG), Color Scale Invariant Feature Transform (C-SIFT), and Maximally Stable Extremal Region (MSER). HOG, C-SIFT, and MSER features were extracted for each leaf, and classifiers were built using the SVM algorithm for HOG and the KD-Tree algorithm for C-SIFT and MSER. These techniques were evaluated on two leaf datasets: a personally built dataset and the Flavia dataset. The experimental results revealed that HOG achieved accuracy of 98% on the personal dataset and 97% on the Flavia dataset. Additionally, C-SIFT demonstrated accuracy of 98% for both datasets, while MSER achieved accuracy of 96% and 90% for the personal and Flavia dataset, respectively. In another study, Saleem *et al*. (2019) presented a plant classification method comprising preprocessing and segmentation, feature extraction, dimensionality reduction, and classification. The leaf was first separated from the background in the preprocessing step using segmentation, followed by the application of a Laplacian operator to smooth leaf edges and regulate small glitches. From the leaf images, shape, texture, and color features were extracted to differentiate between various types of leaves based on their geometry and texture. The feature set comprised 11 shape features, 7 statistical features, 5 venation features, and Fourier descriptors. To reduce the dimensionality of the feature vector, principal component analysis (PCA) was chosen to eliminate redundancy in the feature set dimension. A feature space of ten new features was formed by selecting only the first ten principal components, corresponding to almost 95% of the existing feature space. In the classification stage, the resultant feature vectors were classified into their respective plant species. A variety of classification methods were utilized, including k-nearest neighbor (KNN), decision tree, naïve Bayesian, and multi support vector machine (SVM) classifiers. The proposed method achieved the highest accuracy of 98.93% on the Flavia dataset and 97.75% on the self-collected dataset when KNN was used as the classifier, without dimensionality reduction. Trukoglu and Hanbay (2019) proposed novel methods inspired by the Local Binary Patterns (LBP) method for leaf recognition. Prior to the application of the methods, morphological operations were performed for size reduction and color channel separation. The proposed approaches included filtering operations based on the regional and overall mean, using the R and G color channels in the Overall Mean-LBP (OM-LBP) and Region Mean-LBP (RM-LBP) methods, respectively. The authors also introduced a third method, ROM-LBP, which combined the parameters from OM-LBP and RM-LBP using both color channels. The attribute parameters obtained from these methods were classified using the Extreme Learning Machine method. The performance of the methods was evaluated on four datasets, where ROM-LBP achieved the highest accuracy of 98.94% on Flavia, 83.71% on ICL, and 92.92% on Foliage. Meanwhile, OM-LBP recorded the highest accuracy of 99.46% on the Swedish dataset. In Liu *et al*. (2018), a plant leaf classification method was presented that was based on a ten-layer CNN. The CNN architecture used in this method was based on the LeNet model. The method utilized the feature extraction and classification capabilities of the CNN to achieve accurate leaf classification. Data augmentation techniques, including horizontal flip, vertical flip, noise, color jittering, and rotation, were applied to expand the size of the leaf database. The experimental results showed that the proposed method achieved high overall accuracy of 87.92% on the Flavia dataset. A computerized vision system was developed by Isik et al in 2024, in order to distinguish between two proprietary rice species (Zulkifli*et al*., 2011). A total of 3810 rice grain's images were taken for the two species, processed and feature inferences were made. 7 morphological features were obtained for each grain of rice. With these features, models were created using LR, MLP, SVM, DT, RF, NB and k-NN machine learning techniques and performance measurement values were obtained. Success rates in the classification were obtained 93.02% (LR), 92.86% (MLP), 92.83% (SVM), 92.49% (DT), 92.39% (RF), 91.71% (NB), 88.58% (k-NN). When we look at the results of the success rate of obtain, it is possible to say that the study achieved success. Pallavi *et al* in 2021 developed a CNN based approach to identify the hibiscus species (Pallavi *et al*., 2021). Within that Hibiscus species dataset was passed to CNN algorithm for training and testing to obtain the type of classification of hibiscus leaf species. For which dataset was separated into 2 parts that is 80 of the hibiscus leaf images is passed to CNN model for training with batch size 5 and epochs 10 and it resulted with accuracy value 0.920 and loss value 0.1905. The 20 of the hibiscus leaf images is passed to CNN model for testing with batch size 5 and epochs and it resulted with accuracy value 0.9020 and loss value 0.2017 (Pallavi *et al*., 2021). But to develop a classification system for the entire Malvaecae family was a herculean task. So far no one will carry out any work related to the image processing and plant recognition of Malvaceae family. Therefore, the present work focused on the image processing and plant recognition of Malvaceae family. The main purpose of this study was to assess deep convolutional neural networks with transfer learning for the identification of different plant leaves of Malvaceae family. The machine learning approaches and decision trees were utilized with transfer learning technique to lessen the training time and enhance the functional capabilities of neural networks. CNN was mainly used for plant recognition. The current work uses a 170 leaf image of three plant species in Malvaceae family. The 170 images were used for training purpose. It entire work carried out with the aid of Python, numpy, scikit, and open cvIt was observed that the CNN model performed effectively in distinguishing the plant leaf image and achieved 71 percentage accuracy.

## CONCLUSION

The integration of GLCM, Otsu thresholding, and CNN techniques provides a comprehensive and effective solution for plant classification system. The proposed solution leverages the strengths of each method to achieve accurate and reliable results in identifying and classifying plants based on digital images. By employing this integrated approach, farmers, researchers, and agricultural practitioners can effectively detect and classify plants, leading to timely interventions and optimized taxonomic management strategies. Advancements in image processing techniques, deep learning algorithms, and the availability of comprehensive and diverse datasets will further enhance the accuracy and reliability of plant classification systems. The integration of GLCM, Otsu thresholding, and CNN techniques offer a promising solution for plant classification, providing valuable insights for those who are interested in taxonomy and contributing to the overall community those who are interested in agricultural systems. Implementation of Deep Learning in agriculture has increased enormously in the last decade, including their use to classify different plant leaves. More recently, many deep learning-based approaches for plant leaf stress identification have been proposed in the literature, but there are only a few partial efforts to summarize various contributions. In this study, we train a model with 170 image data to classify leaf types. Based on the experiment result, we obtain an accuracy was 71, Based on experimental results, we can conclude that the implementation of CNN architecture for classifying leaf types can help develop deep learning in agriculture research. We hope that this research can help the public and researchers in recognizing types of plants in Malvaecae family.

## REFERENCES


1. Arafat, S. Y., Saghir, M. I., Ishtiaq, M., & Bashir, U. (2016, July). Comparison of techniques for leaf classification. In 2016 Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP) (pp. 136-141).
2. Cerutti, G., Tougne, L., Mille, J., Vacavant, A., &Coquin, D. (2013). Understanding leaves in natural images–a model-based approach for tree species identification. Computer Vision and Image Understanding, 117(10), 1482-1501.
3. Das, G., Shin, H. S., Ningthoujam, S. S., Talukdar, A. D., Upadhyaya, H., Tundis, R., & Patra, J. K. (2021). Systematics, phytochemistry, biological activities and health promoting effects of the plants from the subfamily bombacoideae (family Malvaceae). Plants, 10(4), 651.
4. Das, U., & Islam, M. S. (2019). A review study on different plants in Malvaceae family and their medicinal uses. Am. J. Biomed. Sci. Res, 3(2), 94-97.
5. De Smedt, T., &Daelemans, W. (2012). Pattern for python. The Journal of Machine Learning Research, 13(1), 2063-2067.
6. ERARSLAN, Z. B., & KOÇYİĞİT, M. (2019). The important taxonomic characteristics of the family Malvaceae and the herbarium specimens in ISTE. Turkish Journal of Bioscience and Collections, 3(1), 1-7.

7. Habtemariam, S. (2019). Medicinal foods as potential therapies for Type-2 diabetes and associated diseases: the chemical and pharmacological basis of their action. Academic Press.

8. Hao, J., & Ho, T. K. (2019). Machine learning made easy: a review of scikit-learn package in python programming language. Journal of Educational and Behavioral Statistics, 44(3), 348-361.

9. Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357 362.

10. Lutz, M. (2001). Programming python. " O'Reilly Media, Inc.".

11. Pallavi, M., Prasad, B. M., Shanthi, P., Reddy, V. L. N., & Kumar, A. N. (2021). Multi trait genotype-ideotype distance index (MGIDI) for early seedling vigour and yield related traits to identify elite lines in rice (Oryza sativa L.). *Electronic Journal of Plant Breeding*, *15*(1), 120-131.

12. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... &Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

13. Saleem, G., Akhtar, M., Ahmed, N., & Qureshi, W. S. (2019). Automated analysis of visual leaf shape features for plant classification. Computers and Electronics in Agriculture, 157, 270-280.

14. Sheikh, H., Prins, C., &Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In Mission AI: The New System Technology (pp. 15-41). Cham: Springer International Publishing.

15. Shetty, P., & Balasubramani, R. (2021). CNN Based Approach to Identify Hibiscus Plant Species. Iconic Research And Engineering Journals, 7(12), 22-26. Retrieved from IRE Journals. (e-ISSN: 2456-8880)

16. Tsopgni, W. D. T., Azebaze, A. G. B., Teinkela, J. E. M., Ndjakou, B. L., Bosyom, F. F., Tchaleu, B. N., &Vardamides, J. C. (2019). New unsaturated fatty acid and other chemical constituents from the roots of Cola rostrata K. Schum.(Malvaceae). Biochemical systematics and ecology, 86, 103913.

17. Turkoglu, M., &Hanbay, D. (2019). Leaf-based plant species recognition based on improved local binary pattern and extreme learning machine. Physica A: Statistical Mechanics and its Applications, 527, 121297.

18. Xu, Z., Deng, M., Xu, Z., & Deng, M. (2017). Malvaceae (pp. 717-735). Springer Netherlands. Zulkifli, Z., Saad, P., &Mohtar, I. A. (2011, December). Plant leaf identification using moment invariants & general regression neural network. In 2011 11th International conference on hybrid intelligent systems (HIS) (pp. 430-435). IEEE.