label of target class is determined by considering the best nearest neighbor obtained with the K-NN classifier based on the Euclidean distance (l_2 norm) measure. If the values of Euclidian distances (dst 1, dst 2, …, dst n) are less than the predetermined associated threshold values (thrsh 1, thrsh 2, …, thrsh n), then the test and train images are accepted as similar to each other according to global features. The threshold value is specified as max distance between the test image and any train image in a processed class. So, the value of threshold is computed for each class.

In test stage, the class label of processed leaf object is determined by utilizing the local features. In case of local features, the minimum Hausdorff distance measure between the curve points of test and train sample of processed class is considered as distinctive factor. For this purpose, the curve points are obtained by performing the Bézier Curves on the segmented image. To assign an incoming leaf object into a unique class, the result returned from K-NN classifier with respect to Hausdorff distance is carried out to determine nearest and similar class. The general working principles of K-NN classifier and Hausdorff distance are explained under the related chapters in more detail.

As a summary, the global features are introduced prior to local features in decision mechanism. The reason why we have used the global and local features together could be expressed as improving the recognition performance and reducing the search space, which is the time to assign a test image into a unique class. With this way, it is believed that a robust system is designed for leaf object recognition. On the other hand, we do not have to launch the classification process between the test and training image as long as they are dissimilar as global.

2.1 Global Features: Gross Shape Features and Central Moments

To improve accuracy and reduce search space, the contributions of different features are extensively evaluated in the experiment stage of global feature extraction. The utilized global features consist of 11 features and are separated into two categories: gross shape features (7) and moment based features (4), respectively. As a part of morphology, the gross features have been obtained by analyzing the form and structure of objects. In the first stage, a total of 7 features, rectangularity, aspect ratio, mean hue, eccentricity, convexity, length of boundary, and arc length of boundary of object, have been experimented on and considered as gross shape features. A brief description for each gross shape features is given as follows:

- Rectangularity: It is computed by fitting a rectangle on the shape area of leaf object depending upon minimum error measure as much as possible.
- Aspect Ratio: This feature refers to the ratio of the longer to smaller length and is provided by fitted rectangularity.
- Mean Hue: It is obtained as finding the average value of intensity of leaf object in hue channel, which is invariant to light intensity variation. As a contribution of this feature, it can distinguish the objects in similar shape form from each other.
- Eccentricity: It is derived by computing the ratio of radii of a generated ellipse on the shape of leaf object.
- Convexity: Similar to the procedure realized in rectangularity, a convex is fitted to each of the leaf object shapes with respect to a minimum error criterion.
- Length of Boundary: It is obtained with a simple way as only calculating the number of pixels stated on boundary.
- Arc Length of Boundary: The arc length feature is the measure for determining the distance along the curved line. Mathematically, the arc length of a parametric form is obtained with the formula $l = \sqrt{(dx/dt)^2 + (dy/dt)^2}$.

In the second stage, for the selection of global features, a total of 4 moment based features have been added to the vectoral feature set. In computer vision, it is widely accepted that image moment is useful for identifying each of the objects in the scene by using weighted average of the image pixels' intensities or considering these moments as a function. Basically, it presents the information about shape of an object by providing its area, orientation, thickness, skewness, and such characteristics. By considering these advantages, in this study, central moments of image are experimented on because of their significant properties being invariant to orientation, size, and translation. The general formula for calculating any central moment is defined as

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (1)$$

where $\bar{x} = \frac{\mu_{10}}{\mu_{00}}$ and $\bar{y} = \frac{\mu_{01}}{\mu_{00}}$ are the components of the centroid, and p and q refer to degree of moments. To obtain the discriminated features, the central moment procedure has applied grayscale of processed image by choosing the values of the $p = 2$ and $q = 2$. With this way, 4 different central moments including μ_{02}, μ_{20}, μ_{12} and μ_{21} are extracted from an input image.

Hence, 11 global features from each object are obtained. For each class, the obtained 11 features have been used to compute the covariance matrix, eigenvalues, and eigenvectors. As emphasized in the study related to the subspace based features extraction strategy (Bilginer Gulmezoglu et al., 1999), since using the eigenvectors corresponding to zeros eigenvalues for projection is a better way to represent the common matrix of a class, in this regard, the obtained 11 global features have been projected by using the eigenvectors corresponding to zeros eigenvalues in order to obtain the projection matrix for a processed class. Thus, the same procedure is repeated until obtaining the projection matrix of each class.

In the test stage, the features of the incoming test object are projected by using the projection matrix of each sample in all classes individually. Also, the global features of the training images are obtained with the same procedure. Once the features returned from the test and train images are obtained, the difference of l_2 norm between vectoral sets of a test and train images of each class is taken into account in terms of finding target class. If the difference of l_2 norm between the vectoral feature sets is greater than a predefined threshold value, then the proposed system automatically makes a decision that the incoming test object does not belong to the processed class, and the procedure continues with the next class. Otherwise, it is assumed that the target class is reached, and the classification procedure should be launched by utilizing the local features. This way, the computation time for classification and search space is reduced. The threshold value is determined through some observations on experimental results.

2.2 Local Features: The Curve Points on Fitted Bézier Curve

In the local feature stage, the curve points depicted on the fitted curve along the boundary of segmented leaf object have been considered as discriminant features. Generally, the curve fitting is the process of designing a function in order to capture the trend in the data across a given image. The objective of a curve fitting methodology is approximating the shape of the control polygon, which consists of some edge segments. Therefore, if the fitted curve results in too big error, the data points would be divided into two different sets. However, this is an undesired situation when considering the curve as a whole. For this purpose, researchers have proposed different types of curve fitting algorithms such as utilizing line segments (Horng and Li, 2001) and polynomial curve fitting (Zhang et al., 2006), algebraic curve fitting (Mizuta, 1996), or other approaches (Kvam and Vidakovic, 2007).

Bézier curves are one of the most frequently used parametric curves in image processing and related fields. They are interpolation functions, which means that the generated values are somewhere between the processed points. There are a number of good reasons why Bézier curves could be chosen on behalf of the curve fitting (Iglesias). These reasons could be summarized as follows:

- Global Control: Moving a control point changes the curve shape completely.
- Local Control: If a control point alters, only a part of the curve is modified.
- Interpolation: End control points are always interpolated.
- Tangency: The tangent of the end points is along the line joining the nearest point of endpoints to endpoints.

- Convex Hull Property: The curve is contained in the convex hull of its defining control points.
- Variation Diminishing Property: No straight line intersects a Bézier curve more times than it intersects its control polygon.
- Subdivision: They can be rendered in many ways.
- Degree Raising: A given Bézier curve of degree n (with control points P_i) can be expressed in terms of a new basis of degree n+1.

A parametric Bézier curve piece of degree n is defined as

$$C(u) = \sum_{k=0}^{n} p_k B_{k,n}(u) \tag{2}$$

where p_k and $B_{k,n}(u)$ are the control points and Bernstein polynomial, respectively.

$$B_{k,n}(u) = \binom{n}{k} u^k (1-u)^{n-k}, \qquad k = 0,\ldots,n \tag{3}$$

where n is the polynomial degree, k is the index, u is a variable, and $\binom{n}{k}$ is binomial coefficient. For cubic Bézier curve, n=3.

In this study, the curve points on Bézier curves are accepted as local discriminative features. The aim under the idea of using local features is improving the robustness of decision system. Once the similarity between target and test image is obtained by using the global features, then the local features would be utilized to realize the matching procedure. Moreover, since the overall structures of curves stated on the leaf objects are dissimilar, the curve points have significant effects and play a vital role in identifying the leaf objects.

2.3 Hausdorff Distance

Hausdorff distance is a measure to determine how two sets of points are "close" to each other. It is named after its inventor, Felix Hausdorff, and has been utilized for some computer vision tasks including image matching, object tracking and classification, computer-vision-enabled ophthalmic augmented reality environment, comparing 2D images of the 3D world, and visual navigation of robots.

Let us suppose that $X = \{x_1, \cdots, x_n\}$ and $Y = \{y_1, \cdots, y_n\}$ are two finite points, then the Hausdorff distance between them is computed by the following formula (Huttenlocher et al., 1993):

$$H(X,Y) = \max(h(X,Y), h(Y,X)) \tag{4}$$

and $$h(X,Y) = \max_{x \in X} \min_{y \in Y} \|x - y\| \tag{5}$$

where $\|.\|$ refers to the underlying norm on the points of X and Y *(e.g. the l_2 norm)*. According to the given reference study, the Hausdorff distance is directed, not symmetric, and reflects the farthest distance from any point of A and B and vice versa. Firstly, the distance between any points of A and each point of B is computed based on a norm measure, then the shortest distance is taken into account. This procedure is repeated for all points on A. Finally, the farthest distance is considered as Hausdorff distance between A and B. Since the Hausdorff distance is not symmetric, the farthest distance from A to B is not equal to that from B to A.

3 DEEP LEARNING FOR LEAF CLASSIFICATION

Recently, Convolutional Neural Network (CNN) methods achieved astonishing success in various classification and image processing tasks when compared to feature engineering techniques. Typically, the learning mechanism of a CNN algorithm relies on searching the discriminative spatial features with massive convolutional filters (feed-forward) and optimizing these filters with a defined backpropagation rule (feed-backward). The raw data mapped generalized and reduced size of hierarchical features with the help of weighting filters. Of the family of CNN architectures, the uncountable designed models have been proposed for task-oriented purpose. For classification-driven tasks, the AlexNet, VGG16, and GoogleNet have been widely used to improve the performance of leaf recognition. To keep pace with literature, we analyzed the performance of another robust CNN method, namely, Inception-v3 model. What makes the Inception-v3 model more attracting is that it gives superior recognition results along with parameter-free testing procedure and reduced number of convolutional filters. The model includes 48 layers, which is deep and more efficient that VGG16 in terms of accuracy. Motivated by the transfer learning, we retrain Inception-v3 by tuning the last head of model according to class numbers of Flavia and Swedish datasets. To restrain the overfitting and underfitting cases in case of training a CNN method, one has to apply the appropriate the data augmentation process with regard to handled data. For this purpose, we performed some methods including shear transformation on shape of image (0.1 ratio), zoom transformation (0.1 ratio), horizontal flip, and random rotation (randomly rotating the image clockwise or anticlockwise with 90 angle). Trainable parameters of the model are about 22.5 million (22, 536, 367), and the disk usage of the model is nearly 265 MegaBytes (MB) for Swedish dataset and 275 MB for Flavia dataset.

4 EXPERIMENTAL STUDY

4.1 Datasets

To make a benchmark evaluation, the proposed system has been conducted on the two well-known leaf datasets, Flavia (Wu et al., 2007) and Swedish (Söderkvist, 2001). The reason why we selected these datasets could be attributed to their popularity in the area of identification plant species by various researchers.

Flavia dataset consists of 32 class of leaf species. Totally, 1907 leaf samples are available as involving about 50 to 70 samples per each class. The sample images are at the 1600x1200 resolution, and the color channel is RGB.

The second dataset is Swedish, which consists of 15 leaf classes. Totally, 1125 samples are included as 75 samples per each species. The samples from the Flavia and Swedish datasets are given in Fig. 2, where the top images refer to Flavia and bottom ones refer to the Swedish dataset. In case of experimental stage with handcrafted features, all images are resized as 320x240 pixels in order to improve the speed of the system. Also, for the Inception-v3 model, each leaf image is resized to 224x224x3 form.

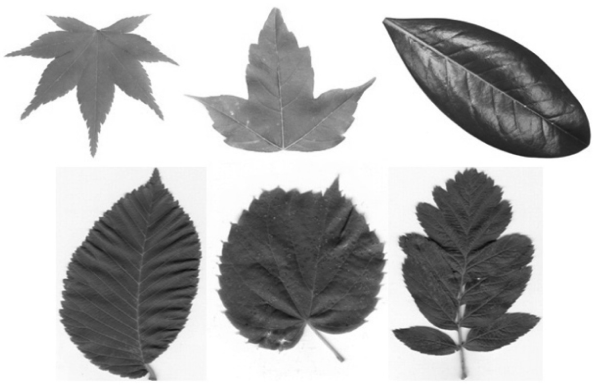

Figure 2. The overview of some samples from Flavia and Swedish datasets.

4.2 Experimental Results

To verify the performance of proposed method, an extensive experimental study is performed using a smart decision system with the global and local features, which are forwarded into two widely used datasets (Flavia and Swedish). As a hybrid approach, using the global and local features for leaf object recognition seems more useful when observing the results of the proposed method in comparison with some recently proposed methods.

To evaluate the performance of the system, the accuracies of the recently published studies on leaf recognition using the same databases are given in terms of accuracy metric. To make a benchmark evaluation, each class of the datasets is split into training and test sets. Similar to the procedures realized in compared works, the holdout method is used for calculating the classification accuracy. This way, only the first $1/3$ images from each class are chosen for training stage, and the rest of the images are remaining for testing purposes in case of simulations on handcrafted features. Although, in some studies, the $k-fold$ cross-validation has been preferred as a model evaluation method, but for this study, we have only preferred the holdout method because of experimenting the large size of datasets and model. For Inception-v3 model, we have considered a random splitting ratio for each class as 70:10:20 for train, validation, and test sets, respectively.

Table 1. Comparison of obtained accuracy results (%) on Flavia dataset.

Performance of Handcrafted Features (%)	
Proposed	**96.78**
SVM-BDT (Singh et al., 2010)	96.00
PNN (Wu et al., 2007)	90.00
Fourier Moment (Singh et al., 2010)	62.00
Performance of Deep Learning Models (%)	
Inception-v3(Proposed)	**98.95**
D-Leaf (Tan et al., 2018)	94.88
LeNet (Liu et al., 2018)	87.92
ResNet50(Modified) (López Barrientos, 2017)	97.00

After experimental simulations on Flavia dataset with handcrafted features, the results of the proposed method and SVM-BDT (Singh et al., 2010), Probabilistic Neural Network (PNN) (Wu et al., 2007), and Fourier Moment (Singh et al., 2010) methods are compared in terms of accuracy rate. Upon inspecting the results given in Table 1, the proposed method gives satisfactory outcomes with 96.78% recognition rate. Moreover, it is obviously seen that the Fourier Moment method, in which the Fourier moments have been performed as a classification technique, shows the worst performance when compared with SVM-BDT and PNN methods. Upon inspecting results of deep learning based models, one can say that there is a competition between performances of each CNN model. To provide the discriminative information of leaf images, an automated system was applied on the basis of CNN approach and named as D-Leaf method (Tan et al., 2018). In the given method, the AlexNet-like structure was applied, and 94.88% score obtained after making experiments with train and test sets, which were split with a ratio of 80% and 20%, respectively. One can emphasize that the referred CNN model is relatively small since it consists of only 3 convolutional layers, 3 max-pooling, and 3 FC layers. In another study (Liu et al., 2018), the LeNet-like architecture was performed on Flavia dataset by using a MATLAB toolbox, namely, MatConvNet. It was reported that the utilized CNN model reached 87.92% overall accuracy rate over 20% of test samples (only 1600 images from 1907). Also, the potential limitation of a modified ResNet50 model was investigated for Flavia leaf dataset (López Barrientos, 2017). Although this model provides the 97% accuracy rate, it suffers from the memory usage and training limitations, since the GPU computing is required for increased number of batch size. Although the scores returned from the proposed method on Flavia dataset are not encouraging as well as the scores for Swedish, this factor could be attributed to the high variations in this dataset and recognition difficulties in larger number classes.

Table 2. Comparison of obtained accuracy results (%) on Swedish dataset.

Handcrafted Features	
Proposed	**94.66**
IDSC (Ling and Jacobs, 2007)	94.13
MDM (Hu et al., 2012)	93.60
FDs (Ling and Jacobs, 2007)	89.60
Deep Learning Models	
Inception-v3(Proposed)	**99.11**
D-Leaf (Tan et al., 2018)	98.09
GoogleNet (Pawara et al., 2017)	98.24
AlexNet (Pawara et al., 2017)	97.81
VGG16 (Zhang et al., 2019)	94.80

To enhance the performance specification, the success of the proposed handcrafted features that were achieved on the Swedish dataset has been compared again with that of some recent methods, which are MDM (Hu et al., 2012), Inner-Distance Shape Context (IDSC) (Ling and Jacobs, 2007), and Fourier Descriptors (FD) (Ling and Jacobs, 2007), which were experimented on the same dataset with different feature extraction procedures. By observing the accuracy results exhibited in Table 2, it is clearly seen that the proposed method outperforms the other aforementioned ones. Additionally, we have addressed the performance of some state-of-the-art deep learning based methods for leaf classification on Swedish dataset. As shown in Table 2, the D-Leaf method achieved the 98.09% discrimination score on Swedish dataset. Moreover, seminal contributions of data augmentation on leaf recognition have been addressed by concentrating on two popular CNN architectures, namely, AlexNet and GoogleNet (Pawara et al., 2017). As shown in Table 2, promising findings were reported as 96.35% and 95.00% for AlexNet and GoogleNet models, which shows that the two models present similar scores when combined with proper data augmentation technique. By motivating from transfer learning concept, the pre-trained VGG16 model was retrained on Swedish dataset (Zhang et al., 2019). The performance of the model accounted for 94.80% accuracy score when using 1/3 of data for training and the rest for testing the system. We can note that the result of Inception-v3 is substantially better than that of the existing models for Swedish dataset. The proposed method achieves a statistically significant improvement, which accounted for the highest score as 99.11% when compared to other popular CNN techniques. One can observe that the existing CNN methods suffer from certain weaknesses in representing the leaf recognition problem with a simple, effective, and efficient model. From the evaluations, it turns out that Inception-v3 model is sufficiently accurate when it comes to building a leaf recognition system with low memory consumption. Although the classification of leaves is a challenging problem because of the high intraspecies variability and low interspecies variation in datasets, all differences in performance of our handcrafted and parameter-free method (Inception-v3 model) are statistically noteworthy in leaf recognition.

4.3 Analysis of Running Time

In this experiment, we have analyzed the running time elapsed per each global and local feature by using the 1907 samples on Flavia dataset. For handcrafted feature extraction, all experiments given in this paper were run on the same hardware (Intel core i5-3210M with 2.5 GHz CPU and 4 GB memory) with software implemented on the Matlab, whereas we have used the python libraries for deep learning model. The running time for each feature is given in seconds (s). By taking the running time shown in Fig. 3 into consideration, it can be concluded that using these features positively affects the performance of system for leaf species classification. Only extracting the points on Bézier Curve keeps a long time as about 2 s. Also, the overall time for features is about 2 s. However, matching the local feature set by using Hausdorff distance is causing the waste of time with a significant ratio.

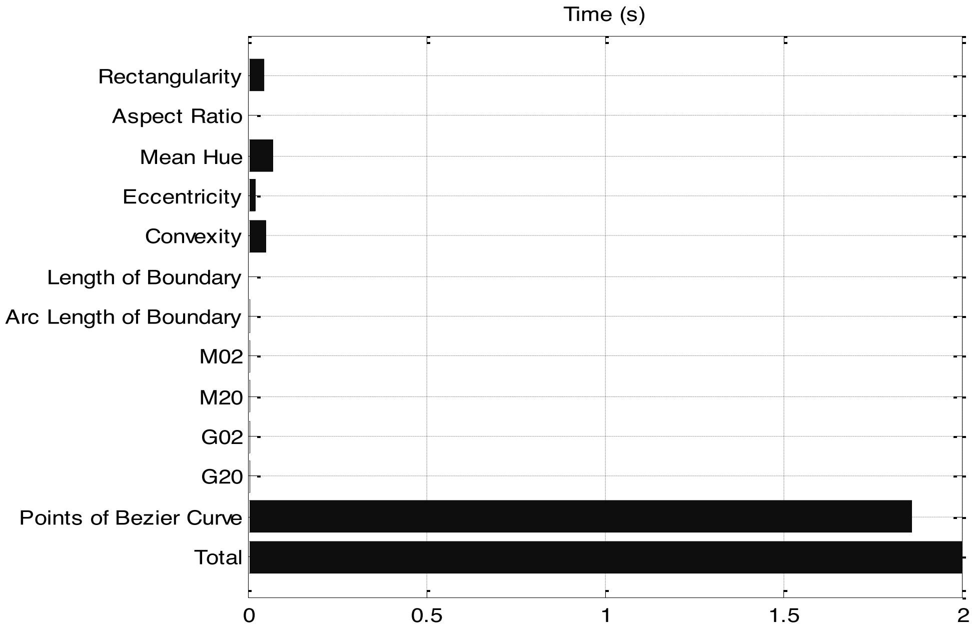

Figure 3. The overall running time for features.

5 CONCLUSION

This paper explains the development of a new approach for the leaf object recognition with a smart decision system. The proposed method consists of two major steps in decision stage. The first step uses 11 global features to determine the target class, which belongs to the leaf object. For this purpose, a vector including different features is constructed, and new feature sets have been generated by employing a projection matrix on this feature vector. Once a new subset of features have been generated, the proposed system automatically makes a decision as whether the incoming test object belongs to the processed class or not by taking the difference of l_2 norms between vectoral sets of a test and train images of each class with respect to a predetermined threshold value. If the target class is specified, then the test object is forwarded to the second stage. At the second stage, the local features are merged as a new feature set and executed to assign the test image in a unique class based on minimum Hausdorff distance criteria.

The implemented system is promising in helping us make an automated classification system for leaf objects in an effortless way and without losing time. The effectiveness of the proposed system has been exhibited through extensive comparisons with other methods using real life public domain data sets with high dimensionality. The experimental results obtained on Flavia and Swedish dataset reveal that the selected features not only are discriminative, but also give strong results as contribution to the classification task in terms of accuracy rate. Moreover, the proposed Inception-v3 model achieves better classification accuracy compared to some recently proposed CNN based methods. The unsatisfying results presented by the compared methods could be attributed to the lack of the stability in the executed feature subsets and the performance of selected classifier in classification stage. As a result, all obtained numerical results prove the robustness of the proposed method.

REFERENCES

Bilginer Gulmezoglu, M., Dzhafarov, V., Keskin, M., Barkana, A., 1999. A novel approach to isolated word recognition. Speech and Audio Processing, IEEE Transactions on **7**, 620-628.

Chaki, J., Parekh, R., 2011. Plant leaf recognition using shape based features and neural network classifiers. International Journal of Advanced Computer Science and Applications (IJACSA) 2.

Franz, E., Gebhardt, E., Unklesbay, K., 1990. Shape description of completely-visible and partially-occluded leaves for identifying plants in digital images. Paper-American Society of Agricultural Engineers.

Horng, J.H., Li, J.T., 2001. A dynamic programming approach for fitting digital planar curves with line segments and circular arcs. Pattern Recognition Letters **22**, 183-197.

Hu, R., Jia, W., Ling, H., Huang, D., 2012. Multiscale distance matrix for fast plant leaf recognition. Image Processing, IEEE Transactions on **21**, 4667-4672.

Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance. Pattern Analysis and Machine Intelligence, IEEE Transactions on **15**, 850-863.

Iglesias, A., Computer-Aided Geometric Design and Computer Graphics: Bezier Curves And Surfaces.

Jobin, A., Nair, M.S., Tatavarti, R., 2012. Plant Identification based on Fractal Refinement Technique (FRT). Procedia Technology **6**, 171-179.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, pp. 1097-1105.

Kvam, P.H., Vidakovic, B., 2007. Curve Fitting Techniques. Nonparametric Statistics with Applications to Science and Engineering, 241-261.

Ling, H., Jacobs, D.W., 2007. Shape classification using the inner-distance. IEEE transactions on pattern analysis machine intelligence **29**, 286-299.

Liu, J., Yang, S., Cheng, Y., Song, Z., 2018. Plant leaf classification based on deep learning, 2018 Chinese Automation Congress (CAC). IEEE, pp. 3165-3169.

López Barrientos, J., 2017. Leaf Recognition with Deep Learning and Keras using GPU computing.

Mizuta, M., 1996. Algebraic curve fitting for multidimensional data with exact squares distance, Systems, Man, and Cybernetics, 1996., IEEE International Conference on. IEEE, pp. 516-521.

Pawara, P., Okafor, E., Schomaker, L., Wiering, M., 2017. Data augmentation for plant classification, International Conference on Advanced Concepts for Intelligent Vision Systems. Springer, pp. 615-626.

Remagnino, P., Mayo, S., Wilkin, P., Cope, J., Kirkup, D., 2017. Machine Learning for Plant Leaf Analysis, Computational Botany. Springer, pp. 57-79.

Rhouma, M.B.H., Žunić, J., Younis, M.C.J.C., agriculture, e.i., 2017. Moment invariants for multi-component shapes with applications to leaf classification. **142**, 326-337.

Shipunov, A.B., Bateman, R.M., 2005. Geometric morphometrics as a tool for understanding Dactylorhiza (Orchidaceae) diversity in European Russia. Biological Journal of the Linnean Society **85**, 1-12.

Simonyan, K., Zisserman, A.J.a.p.a., 2014. Very deep convolutional networks for large-scale image recognition.

Singh, K., Gupta, I., Gupta, S., 2010. Svm-bdt pnn and fourier moment technique for classification of leaf shape. International Journal of Signal Processing, Image Processing and Pattern Recognition **3**, 67-78.

Söderkvist, O., 2001. Computer vision classification of leaves from swedish trees.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826.

Tan, J.W., Chang, S.-W., Kareem, S.B.A., Yap, H.J., Yong, K.-T., 2018. Deep learning for plant species classification using leaf vein morphometric. IEEE/ACM transactions on computational biology and bioinformatics.

Wang, X.-F., Huang, D.-S., Du, J.-X., Xu, H., Heutte, L., 2008. Classification of plant leaf images with complicated background. Applied mathematics and computation **205**, 916-926.

Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.X., Chang, Y.F., Xiang, Q.L., 2007. A leaf recognition algorithm for plant classification using probabilistic neural network, Signal Processing and Information Technology, 2007 IEEE International Symposium on. IEEE, pp. 11-16.

Zhang, J., Xu, G., Chen, H., Zhao, Q., Yan, W., 2006. Using polynomial curve fitting method to improve image quality in EIT. IEEE E MBC, 6769-6772.

Zhang, S., Wang, H., Huang, W.J.C.c., 2017. Two-stage plant species recognition by local mean clustering and Weighted sparse representation classification. **20**, 1517-1525.

Zhang, X., Zhao, W., Luo, H., Chen, L., Peng, J., Fan, J., 2019. Plant recognition via leaf shape and margin features. Multimedia Tools Applications **78**, 27463-27489.