Journal of Agricultural Informatics (ISSN 2061-862X) 2015 Vol. 6, No. 2:32-44

Hungarian Association of Agricultural Informatics
European Federation for Information Technology in Agriculture, Food and the Environment

Journal of Agricultural Informatics. 2015 Vol. 6, No. 2
journal.magisz.org

A computer-based vision systems for automatic identification of plant species using kNN and genetic PCA

Oluleye Babatunde[1], Liesa Armstrong[2], Jinsong Leng[3], Dean Diepeveen[4]

I N F O
Received 20 Apr 2015
Accepted 23 Apr 2015
Available on-line 29 Jun 2015
Responsible Editor: M. Herdon

Keywords:
Precision Agriculture, Plant Species Classification, Image Processing, Principal Component Analysis, k Nearest Neighbor, Genetic Algorithm.

A B S T R A C T

Precision farming involves integration of different areas of disciplines to lower production costs and improve productivity. One major arm of precision farming or agriculture is the development of computer-based vision systems for automatic identification of plant species. This work involves application of k Nearest Neighbour (kNN) and genetic principal component analysis (GA-PCA) for the development of computer-based vision systems for automatic identification of plant species. As the first contribution, several image descriptors were extracted from the images of plants found in the Flavia dataset. Lots of these image features are affine maps and amalgamation of such massive features in one study is a novel idea. These descriptors are Zernike Moments (ZM), Fourier Descriptors (FDs), Lengendre Moments (LM) Hu 7 Moments, Texture, Geometrical properties and colour features. The GA-PCA (1907 x 41) feature space improved the classification accuracy of kNN from 84.98% to 88.75%.

1. Introduction

Plant species identification is traditionally carried out by manual matching of the plant's features, relating to components of the plant, such as leaves, flowers, and bark, against an atlas (Meeta, 2012) and (Abdul, Lukito, Adhi and Santosa, 2012). According a survey (Babatunde, Armstrong, Leng and Diepeveen, 2015) attempts to automate this process have been made, using features of plants extracted from images as input parameters to various classifier systems. This work further examines the application of genetically selected principal components combined with k Nearest Neighbor (kNN) algorithm. The main strength of this article is the application of a GA to automatically select the minimum number of principal components needed for optimal accuracy based on the available dataset.

2. Literature Reviews

Several plant species recognition systems have been developed based on various features and classifiers. Many of these works were based on artificial neural networks (ANN) as Machine Learning Models due to their adaptability and scalability. Table 1 below shows some recent works on computer-based vision systems for automatic identification of plant species. A report on their weaknesses are shown in the last column. A deeper literature reviews on recent works on computer-based vision

[1] Oluleye Babatunde
School of Computer and Security Science, Edith Cowan University, Perth, WA, Australia;
Department of Information and Communication Technology, Osun State University, Osogbo, Osun State, Nigeria
hezecomp@yahoo.com
[2] Liesa Armstrong
School of Computer and Security Science, Edith Cowan University, Perth, WA, Australia
l.armstrong@ecu.edu.au
[3] Jinsong Leng
Security Research Institute, Edith Cowan University, Perth, WA, Australia
j.leng@ecu.edu.au
[4] Dean Diepeveen
Department of Food and Agriculture, South Perth, WA, Australia
dean.diepeveen@agric.wa.gov.au

systems for plant species identification can be found in the paper (Babatunde, Armstrong, Leng & Diepeveen, 2015).

Table 1. Some existing and recent works on plant recognition systems

Author	**Techniques**	**Weaknesses (comments)**
Zalikha (et al, 2011)	Image Pre-Processing, Moment Invariants, General Regression Neural Network (GRNN).	This work needs optimization of GRNN and more features
Wu (et al, 2007)	Probabilistic Neural Network (PNN), Image pre-processing, Principal Component Analysis (PCA).	The features used are not enough to ensure improved accuracy across large dataset. Moments and colour features should be included. The parameter of PNN also needs to be optimized.
Panagiotis, T. (2005).	Fuzzy Logic Selection, Neural Networks, image pre-processing, principal component analysis.	More features are needed to improve this work.
Valliammal, N., & Geethalakshmi, S. N. (2011)	Fuzzy Segmentation, image preprocessing, wavelet transformation, leaf image moments	More features needed.
Jyotismita, C., & Ranjan, P. (2011	Thresholding method, H-Maxima transformation, Moment-invariants, Centroid-Radii and Neural Networks classifiers	More features and optimization of the ANN classifier used is needed to improve this work.

3. Methodology

This section discusses the adopted methodology in designing the proposed model (kNN-GA-PCA) for automatic identification of leaves. The used dataset is detailed in section 3.1, while section 3.2 discusses Principal Component Analysis (PCA) which was used for both feature transformation and dimensionality reduction. Genetic Algorithm and k Nearest Neighbour are detailed in section 3.3 and 3.4 respectively. The proposed approach includes image acquisition, image pre-processing, image segmentation, feature extraction, automatic selection of number of principal components by GA and image classification.

3.1 Data set

The source of images of leaves used in this study is images of leaves found in the Flavia dataset which is publicly available (Wu et al., 2007). The Flavia dataset is a constrained set of leaf images taken against a white background and without any stem present. The species in the dataset have a varying number of instances as shown (Babatunde, Armstrong, Leng & Diepeveen, 2014a, 2014b). The dataset has 1907 images of 32 species of plants. The proportion of classes in the Flavia dataset is shown in Figure 1. The complete feature space for this work comprises ZMs, FDs, Lengendre Moments, Hu 7 Moments, Texture, Geometrical properties and colour features which are extracted from the Flavia dataset as shown in Table 2 below.

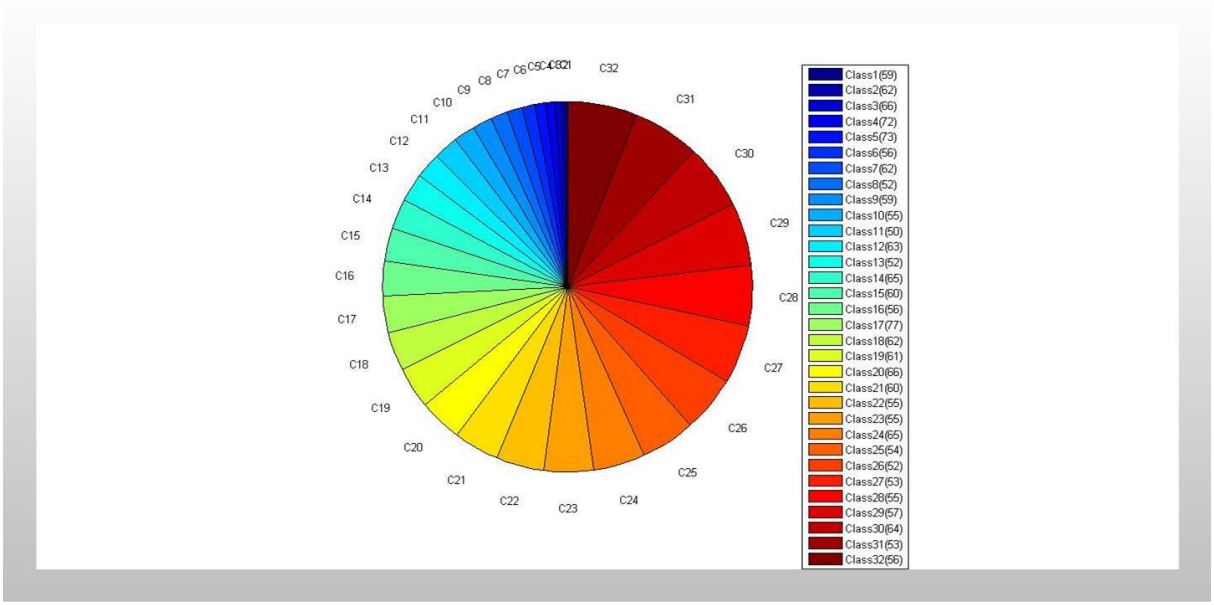

Figure 1. Proportion of plant species in the Flavia dataset

Table 2. 112 Features in the extract from the Flavia dataset

SN	Descriptor	Feature Index	Descriptor Cardinality
1	Zernike Moments (ZM)	F_{01}, F_{02}, …, F_{20}	20
2	Legendre Moments(LM)	F_{21}, F_{22}, …, F_{40}	20
3	Hu 7 Moments (Hu7M)	F_{41}, F_{42}, …, F_{47}	7
4	Texture Features (TF)	F_{48}, F_{49}, …, F_{69}	22
5	Geometric Features (GF)	F_{70}, F_{71}, …, F_{79}	10
6	Fourier Descriptors (FD)	F_{80}, F_{81},…, F_{100}	21
7	Colour features (CF)	F_{101}, F_{102},…F_{112}	12

3.2 Principal Component Analysis (PCA)

PCA is a mathematical procedure (orthogonal transformation from applied linear algebra) that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called the **principal components**. PCA is a dimensionality reduction technique and is useful for dimension reduction when the transformed features have a descriptive power more easily ordered than the original features. It's used in selecting a subset of variables from a large dataset, based on which the original variables have the highest correlations with the principal component. In other words, PCA seeks a linear combination of variables so that the maximum variance can be extracted from the variables. In terms of geometry, PCA is a rotation of the axes of the original variable coordinate system to new orthogonal axes, called principal axes, so that the new axes coincide with directions of maximum variation of the original observations. The property of the maximum variation of the projected points defines the first principal axis and it's the line or direction with maximum variation of the projected values of the original data points. These projected values are called principal component scores (Cambell & Atchlev, 1981). It should be noted that each principal component is a linear combination of the original variables and all the principal components are orthogonal to each other, so there is no redundant information. Thus, PCA is both feature transformation and reduction

technique. The PCA accepts a dataset and rotates it in such a way that the maximum variability can be visible. The operation of PCA is given as follows:

1. **Given a dataset** $\{x_n\}, n = 1(1)N$, and x_n is a N-dimensional vector, (N = 112 for this study)
2. **Task** ?: To project the given data onto an M-dimensional subspace, where M < N , and M, N are positive integers.
3. **Assumption**: The projection is assumed to be represented as

$$y = Ax \quad \text{.....(1)}$$

where

$$A = [u_1^T, u_2^T, u_3^T, ..., u_M^T] \quad \text{.....(2)}$$

and

$$u_i^T u_i = 1 \text{ for } i = 1(1)M \quad \text{.....(3)}$$

The objective now is to maximize the variance of y_n, which is the trace of the covariance of matrix H_y of y_n. Therefore the actual objective now is to find the **maximum of TRACE** (H_y) where

$$H_y = \frac{1}{N}\sum_{n=1}^{N}(y_n - \bar{y})^T \quad \text{.....(4)}$$

and

$$\bar{y} = \frac{1}{N}\sum_{n=1}^{N} y_n \quad \text{.....(5)}$$

If we assume H_x to be the covariance matrix of x_n and since

$$trace\ (H_y) = trace\ (AH_xA^T) \quad \text{.....(6)}$$

the Langrangian multiplier and derivatives gives

$$H_x u_i = \lambda_i u_i \quad \text{.....(7)}$$

where u_i is the largest vector of H_x which corresponds to the ith largest eigenvalue. The first 33 instances of the first 14 principal components of the features derived from the Flavia dataset are given Figure 2. A 2D and 3D plot which allows you to visualize the absolute value and sign of each variable's contribution to the first two or three principal components, and how each observation is represented in terms of components shown in Figures 3, 4 & 5.

	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6	PCA7	PCA8	PCA9	PCA10	PCA11	PCA12	PCA13	PCA14
1	0.0220	0.0062	0.0091	0.0034	0.0354	-0.0241	-0.0027	-0.0085	-0.0146	0.0229	-0.0096	-0.0032	-0.0128	0.0259
2	0.0163	0.0511	0.1338	0.1000	0.0589	0.1984	0.0242	-0.0043	0.0332	0.0425	0.0034	-0.0518	0.0185	-0.0329
3	0.0153	0.0028	0.0148	0.0028	-0.0086	0.0024	-0.0310	0.0772	0.0217	0.1103	-0.0030	-0.0734	-0.0145	0.0124
4	0.0199	0.0033	0.0178	-0.0064	0.0284	-0.0082	0.0012	0.0049	0.0133	0.0237	0.0020	-0.0358	0.0208	-0.0037
5	0.0143	-0.0080	0.0360	-0.0151	0.0189	-0.0123	0.0021	0.0079	0.0111	0.0145	0.0101	-0.0308	0.0279	-0.0024
6	0.0199	0.0014	-0.0041	-0.0073	-0.0067	0.0138	-0.0293	0.0932	0.0136	0.0644	-0.0084	-0.0437	-0.0299	0.0419
7	0.0191	-0.0013	0.0190	-0.0056	0.0292	-0.0185	4.2073e-04	-0.0048	-0.0175	0.0068	-0.0142	0.0015	-0.0209	0.0322
8	0.0200	0.0027	0.0185	-0.0055	0.0314	-0.0111	0.0034	-0.0072	-0.0194	0.0113	-8.1569e-...	0.0019	-0.0176	0.0304
9	0.0197	-0.0018	0.0137	-0.0100	0.0285	-0.0130	7.0806e-04	0.0022	-2.9923e-...	0.0144	0.0072	-0.0149	0.0121	0.0028
10	0.0173	-0.0153	0.0127	-0.0144	0.0190	-0.0057	-0.0118	0.0511	0.0720	0.0129	-0.0063	-0.0243	0.0346	-0.0077
11	0.0164	-0.0105	0.0254	-0.0225	0.0208	-0.0048	0.0065	0.0026	-0.0100	-0.0060	0.0079	8.5887e-04	-0.0119	0.0283
12	0.0187	-0.0074	0.0126	-0.0273	0.0266	-0.0061	0.0093	-0.0042	-0.0163	0.0062	0.0241	0.0042	-0.0153	0.0292
13	0.0180	-0.0073	0.0156	-0.0056	0.0347	-0.0050	-0.0054	0.0196	0.0216	0.0068	0.0042	0.1021	0.0316	0.0029
14	0.0182	-0.0018	0.0232	-0.0063	0.0281	-0.0137	0.0017	-0.0045	-0.0149	0.0078	-0.0087	0.0057	-0.0134	0.0260
15	0.0201	0.0137	0.0067	-0.0094	-0.0134	0.0059	-0.0265	0.0985	0.0099	0.0498	-0.0363	-0.0320	-0.0274	0.0517
16	0.0217	-0.0095	-7.4829e-...	-0.0049	0.0392	-2.9437e-...	3.8089e-04	-0.0098	-0.0177	0.0147	0.0077	-1.9305e-...	-0.0242	0.0301
17	0.0130	-0.0089	0.0208	-0.0038	-0.0140	-0.0026	-0.0337	0.0917	0.0157	0.0695	-0.0111	-0.0501	-0.0317	0.0413
18	0.0182	0.0015	0.0233	-0.0052	0.0348	-0.0059	-0.0021	0.0190	0.0250	0.0156	-0.0019	0.0862	0.0039	0.0244
19	0.0208	-8.7920e-...	-0.0118	0.0029	-0.0027	0.0060	-0.0357	0.0896	0.0140	0.0721	-0.0129	-0.0507	-0.0272	0.0364
20	0.0163	-0.0045	0.0062	-0.0021	-0.0102	-0.0060	-0.0353	0.0902	0.0145	0.0729	-0.0216	-0.0526	-0.0285	0.0377
21	0.0163	0.0047	0.0226	0.0025	-0.0153	0.0079	-0.0305	0.0982	0.0140	0.0466	0.0106	-0.0269	-0.0294	0.0526
22	0.0175	-0.0030	0.0229	-0.0073	0.0322	-0.0045	-0.0023	0.0182	0.0178	0.0040	0.0079	0.1156	0.0212	0.0089
23	0.0186	-6.1982e-...	0.0229	-0.0017	0.0320	-0.0127	0.0015	-0.0081	-0.0143	0.0187	0.0211	-7.1950e-...	-0.0204	0.0266
24	0.0179	-0.0048	0.0212	-0.0072	0.0304	-0.0149	0.0014	-0.0074	-0.0169	0.0146	-0.0023	0.0018	-0.0199	0.0280
25	0.0188	-0.0033	0.0175	-0.0078	0.0310	-0.0170	0.0011	-0.0074	-0.0183	0.0141	-0.0150	0.0021	-0.0198	0.0279
26	0.0178	-0.0097	0.0187	-0.0046	0.0313	-0.0116	-9.6612e-...	-0.0055	-0.0142	0.0106	-0.0069	-7.5025e-...	-0.0168	0.0294
27	0.0180	-0.0043	0.0259	-0.0055	0.0270	1.2694e-04	0.0030	3.0358e-04	-0.0122	-0.0028	0.0068	0.0035	-0.0115	0.0278
28	0.0202	-0.0024	0.0079	-0.0196	0.0332	-0.0092	0.0070	-0.0101	-0.0177	0.0234	0.0024	0.0010	-0.0189	0.0264
29	0.0191	0.0013	0.0223	-0.0030	0.0299	-0.0171	0.0018	-0.0078	-0.0195	0.0106	0.0119	0.0018	-0.0231	0.0332
30	0.0202	-0.0027	0.0126	-0.0076	0.0321	-0.0096	0.0022	-0.0064	-0.0186	0.0077	-0.0072	0.0013	-0.0227	0.0348
31	0.0195	-0.0047	0.0065	-0.0069	0.0368	-0.0122	-0.0066	0.0166	0.0165	0.0139	-0.0111	0.1041	0.0365	-0.0061
32	0.0196	-0.0040	0.0094	-0.0110	0.0356	-0.0047	-0.0028	0.0187	0.0200	0.0087	-0.0071	0.1007	0.0196	0.0187
33	0.0182	-0.0028	0.0208	-0.0071	0.0338	-0.0017	-0.0022	0.0197	0.0228	0.0063	-0.0081	0.0979	0.0057	0.0250

Figure 2. The first 14 principal component of the original feature set

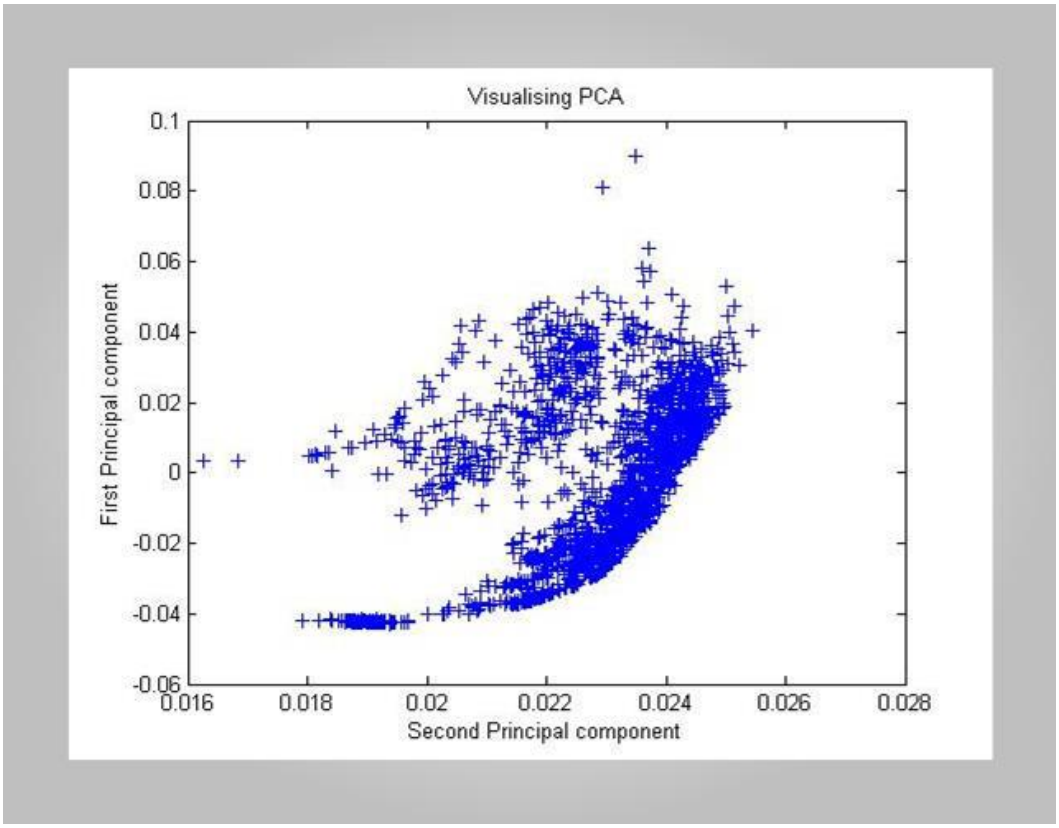

Figure 3. Visualization of two PCA axes

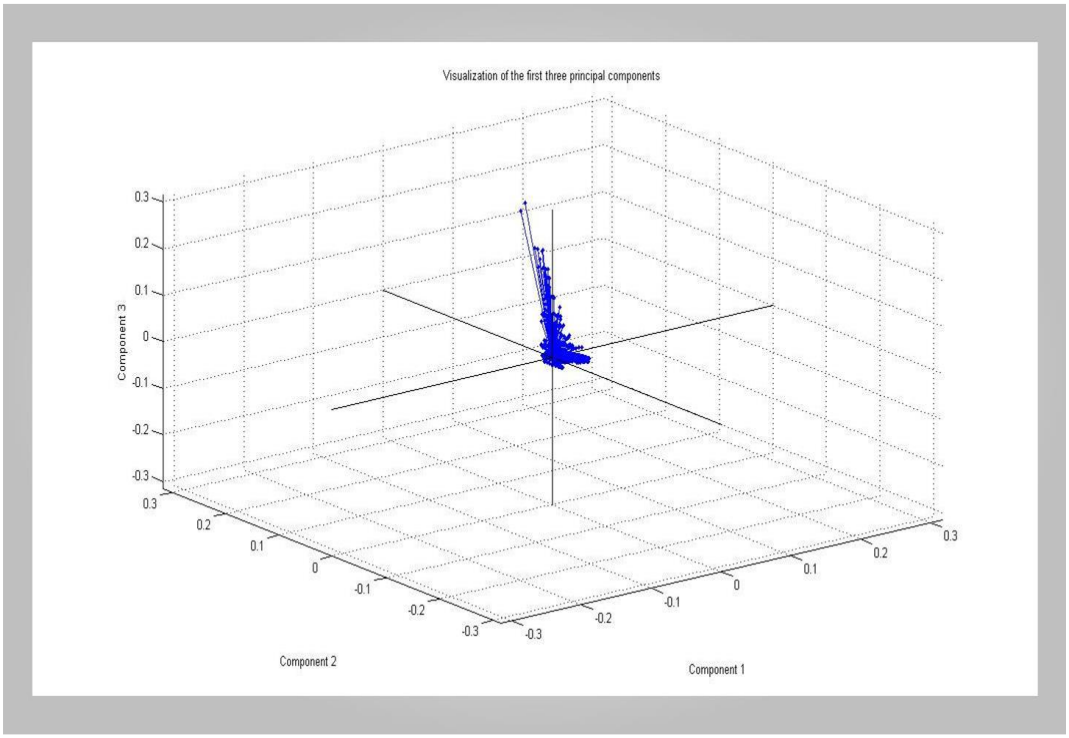

Figure 4. A 3D view of the first three principal components

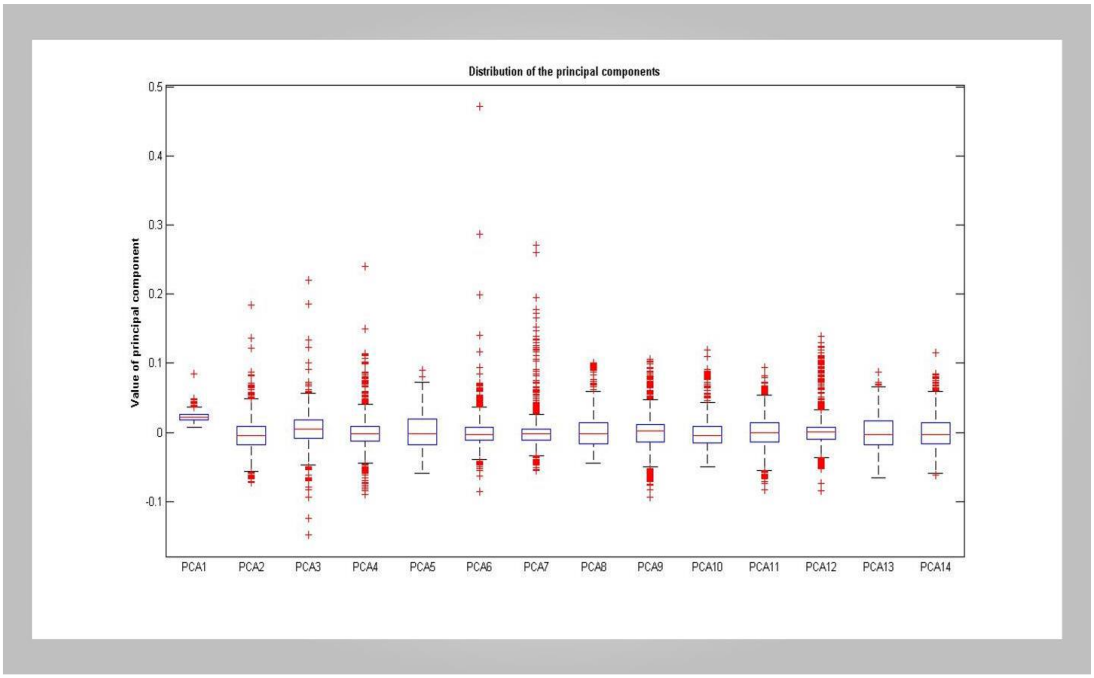

Figure 5. Distribution of the principal components

3.3 Genetic Algorithm (GA)

Genetic Algorithms (GA) can be defined as population-based and heuristic algorithmic searching methods that mimic natural evolution process of man (Melanie, 1999; Tian, Hu, Ma & Ha, 2012; Babatunde, Armstrong, Leng & Diepeveen, 2014d, 2014a, 2014b). GA iteratively employ the use of one population of chromosomes (solution candidates) to get a new population using a method of natural selection combined with genetic functions such as crossover and mutation (in the similitude of Charles Darwin evolution principle of reproduction, genetic recombination, and the survival of the

fittest). In comparative terminology to human genetics, chromosomes are the bit strings, gene is the feature, allele is the feature value, locus is the bit position, genotype is the encoded string, and phenotype is the decoded genotype (Sivanandam & Deepa, 2008). The fitness of each chromosome is evaluated using a function commonly referred to objective function or fitness function. In other words, the fitness function (objective function) reports numerical values which are used in ranking the chromosomes in the population. The fitness function used for both GA and PSO is given as Equation 2. The detailed description of the GA can be found in the companion paper of Babatunde, Armstrong, Leng and Diepeveen (2014b).

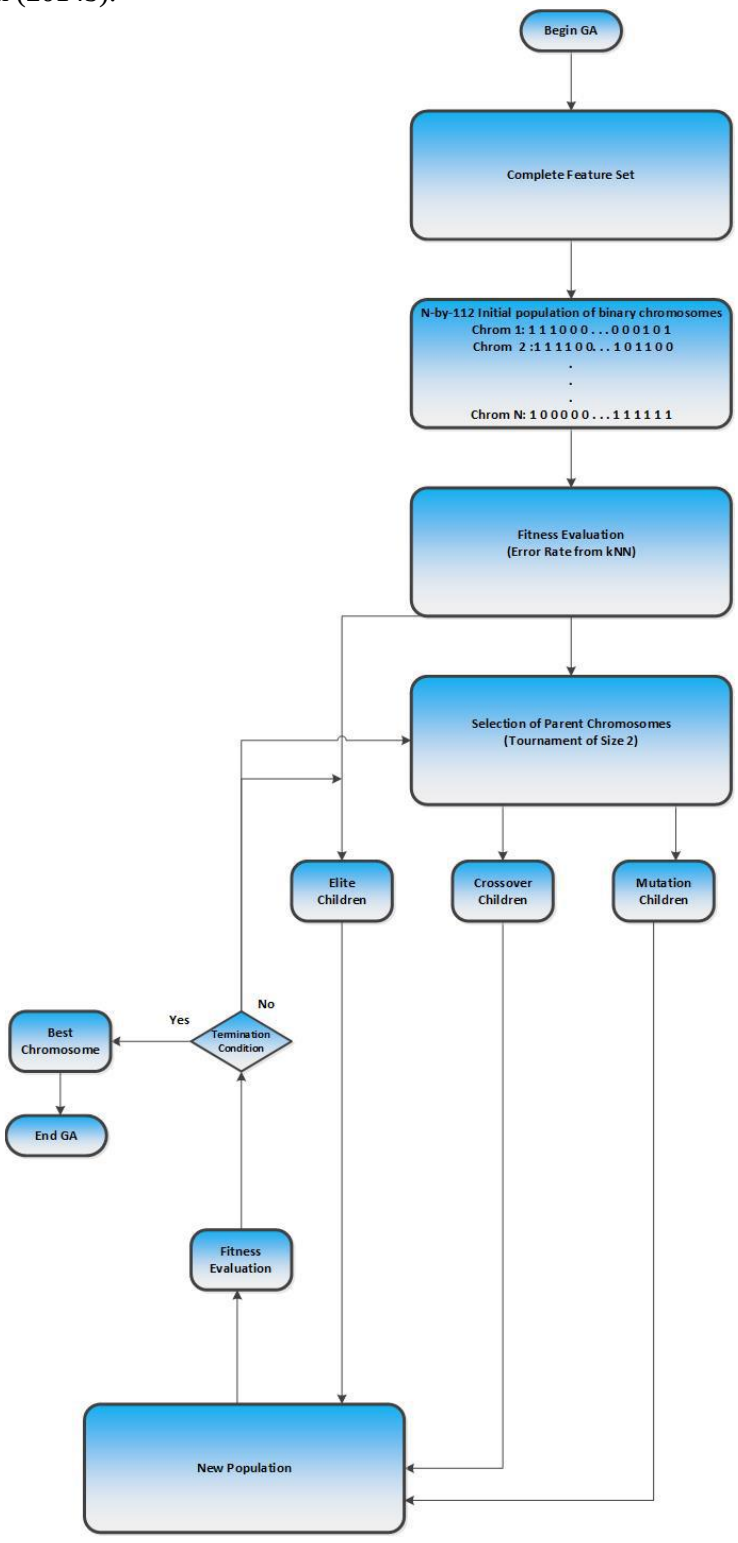

Figure 6. GA-Based Feature Selection ((Babatunde, Armstrong, Leng and Diepeveen, 2014b))

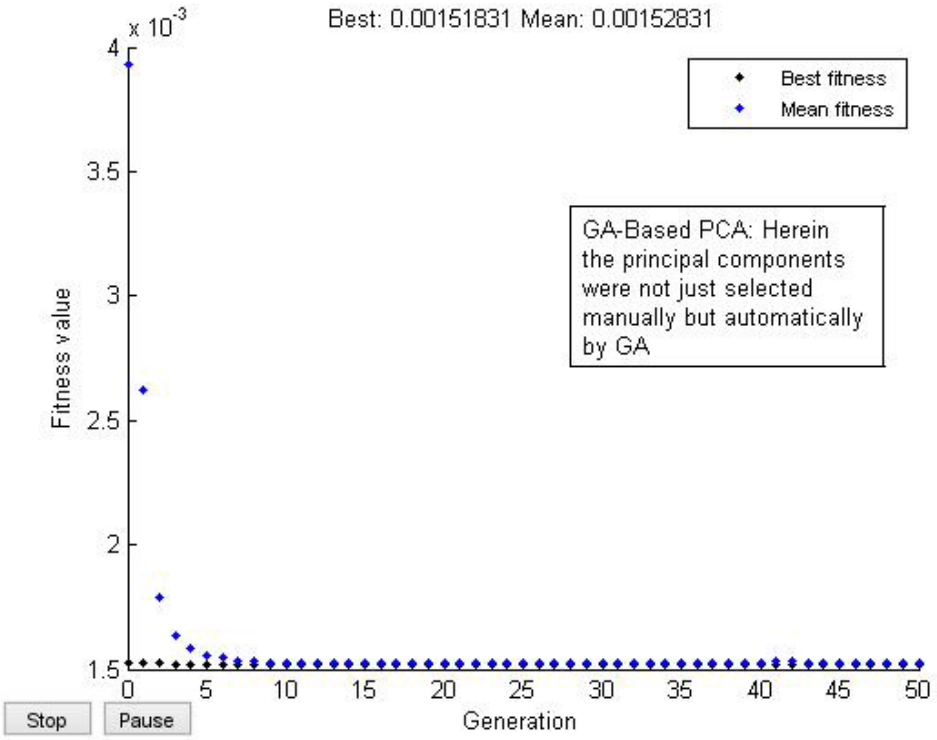

Figure 7. Simulation diagram on genetically selected principal components

3.4 K Nearest Neighbour

The kNN is metric-based algorithm that solves classification problem by looking for the shortest distance between the test data and training sets in the feature space. The distance is generally computed in Pythagorean sense (by finding the square root of the sum of differences). Suppose the training set, using the features in Table 2 is defined as

$$x = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \mathrm{K} & x_{1N} \\ x_{21} & x_{22} & x_{23} & \mathrm{K} & x_{2N} \\ x_{31} & x_{32} & x_{33} & \mathrm{K} & x_{3N} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ x_{M1} & x_{M2} & x_{M3} & \mathrm{K} & x_{MN} \end{bmatrix} \quad \text{.....(8)}$$

where M = 1907, the number of observations in these dataset. The number of features here is N. The kNN algorithm computes Euclidean distance between test data x_{TEST} and all entries in the training sets and then finds the nearest point (shortest distance) from the training set to the test set as:

$$D(x_{TEST}, x_m) = \sqrt{\sum_{m=1}^{M} (x_{TEST} - x_m)^2} \quad \text{.....(9),}$$

where $m = 1,2,3...,M$. The kNN considers only the k nearest neighbours denoted as $\{x_1, \mathrm{K}\ x_k\}$ as the member(s) of the set (a normed linear space).

$$kNNSpace = \{x_j \mid d(x, x_i) \le d(x, x_j)\} \quad \text{.....(10)}$$

The kNN rules involves classifying a test sample, say, x, by assigning it to the most frequently represented among the k nearest samples. The diagram in Figure 8 taken from Mathworks (2013) illustrates 3 Nearest Neighbours as they are the three shortest distances reported. A similar figure generated from this study, showing first 18 neighbours of a test sample from the Flavia dataset is

shown in Figure 9. The kNN counts each category m in the class information and the report classification results based on the expression

$$\arg\max(count(x_m)) \quad \text{.....(11)}$$

subject to

$$\sum_{i=1}^{M} count(x_m) = class \quad \text{.....(12)}$$

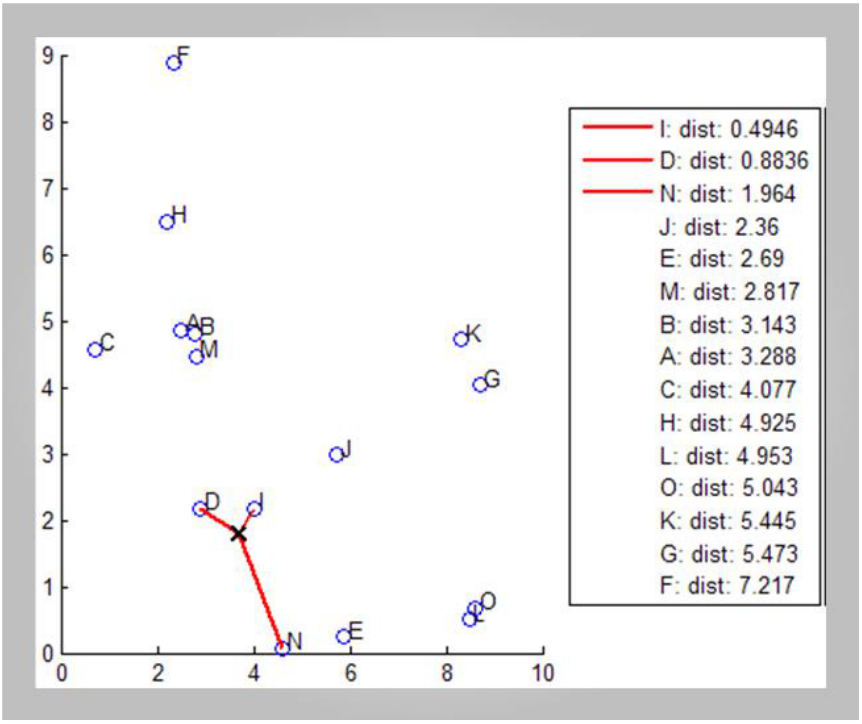

Figure 8. A diagram showing k = 3 nearest neighbours

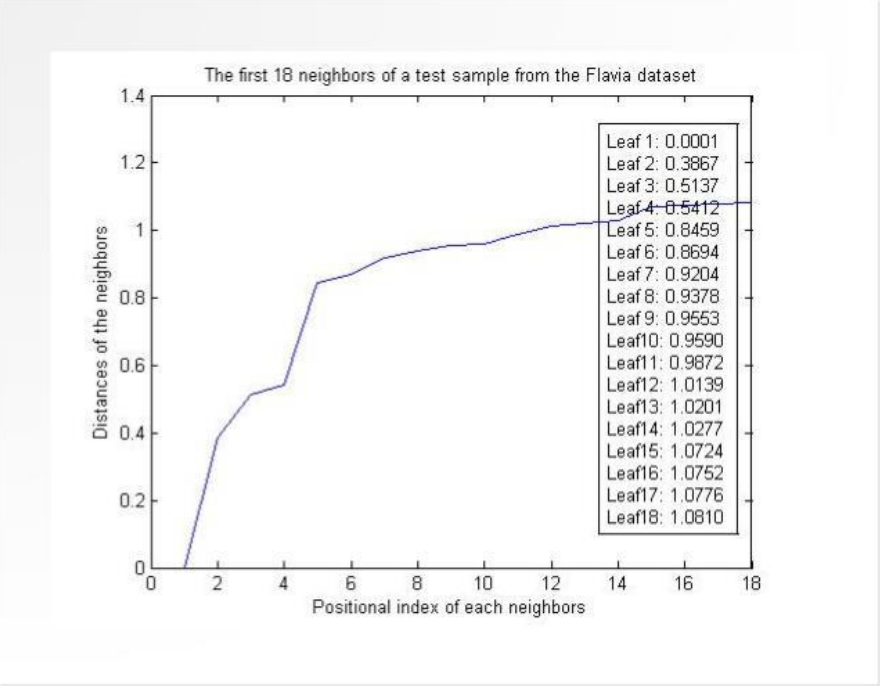

Figure 9. A diagram showing k = 18 nearest neighbours

4. Experimental Design and validation

The research approach used in this work is shown in Figure 10. The classification model was kNN. The different distance metric that can be used with the kNN are Euclidean, standard Euclidean, Mahalanobis, Minkowski, Chebychev, Cosine, Correlation, Jaccard and Spearman distance. The choice distance metric for this study is Euclidean distance. A sample screen shot of the kNN-based classification model. The feature space generated from the Flavia dataset discussed in section 3 was partitioned into two disjoints set (training and test set) via 10-fold cross validation as shown in Figure 11. The feature space (PCA space) itself was a reduced feature space as GA was used to automatically select the number of principal components (PC) finally used. The number of PCs was 41. The implementation was done in MATLAB 2013.

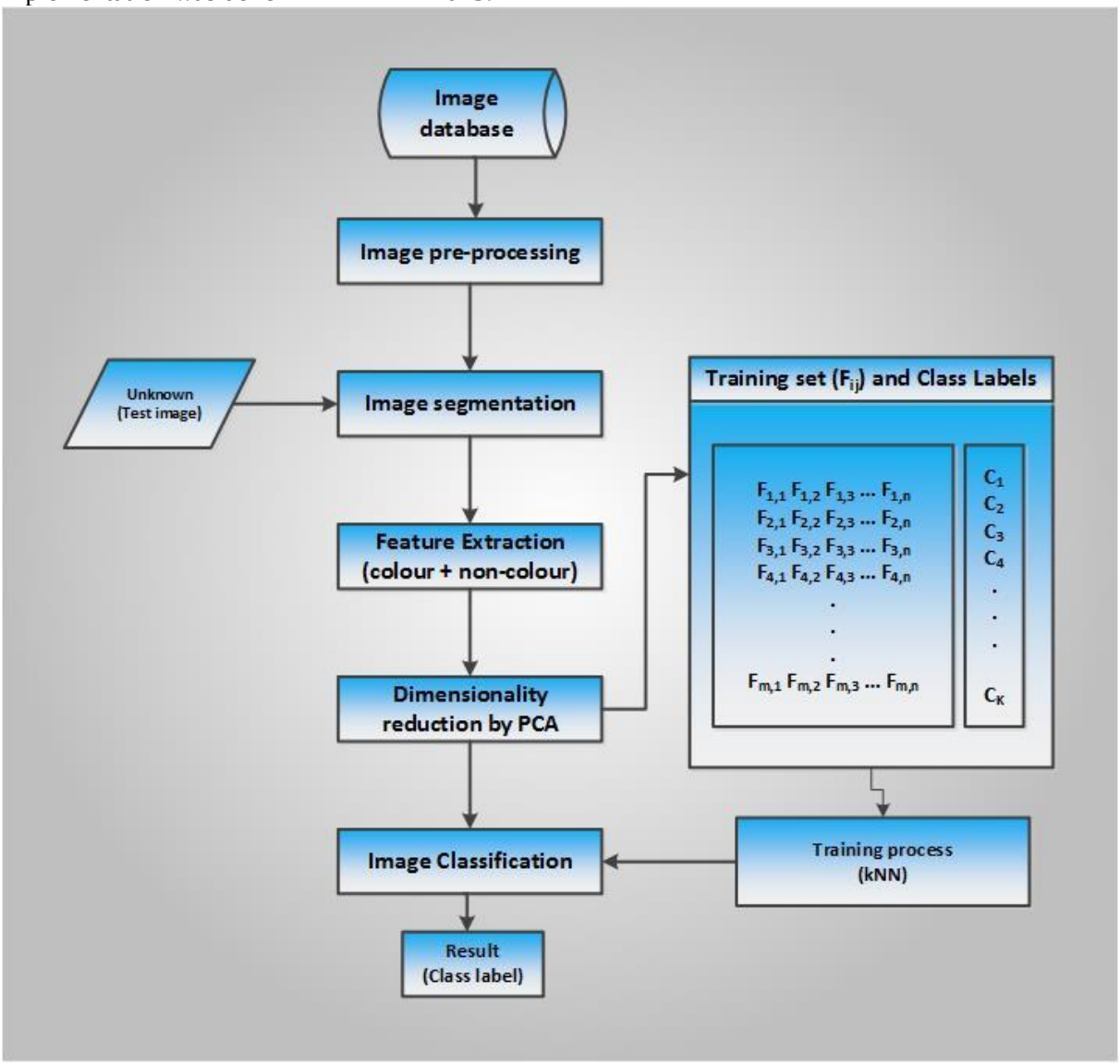

Figure 10. A research approach on computer-based vision system for automatic identification of plant species.

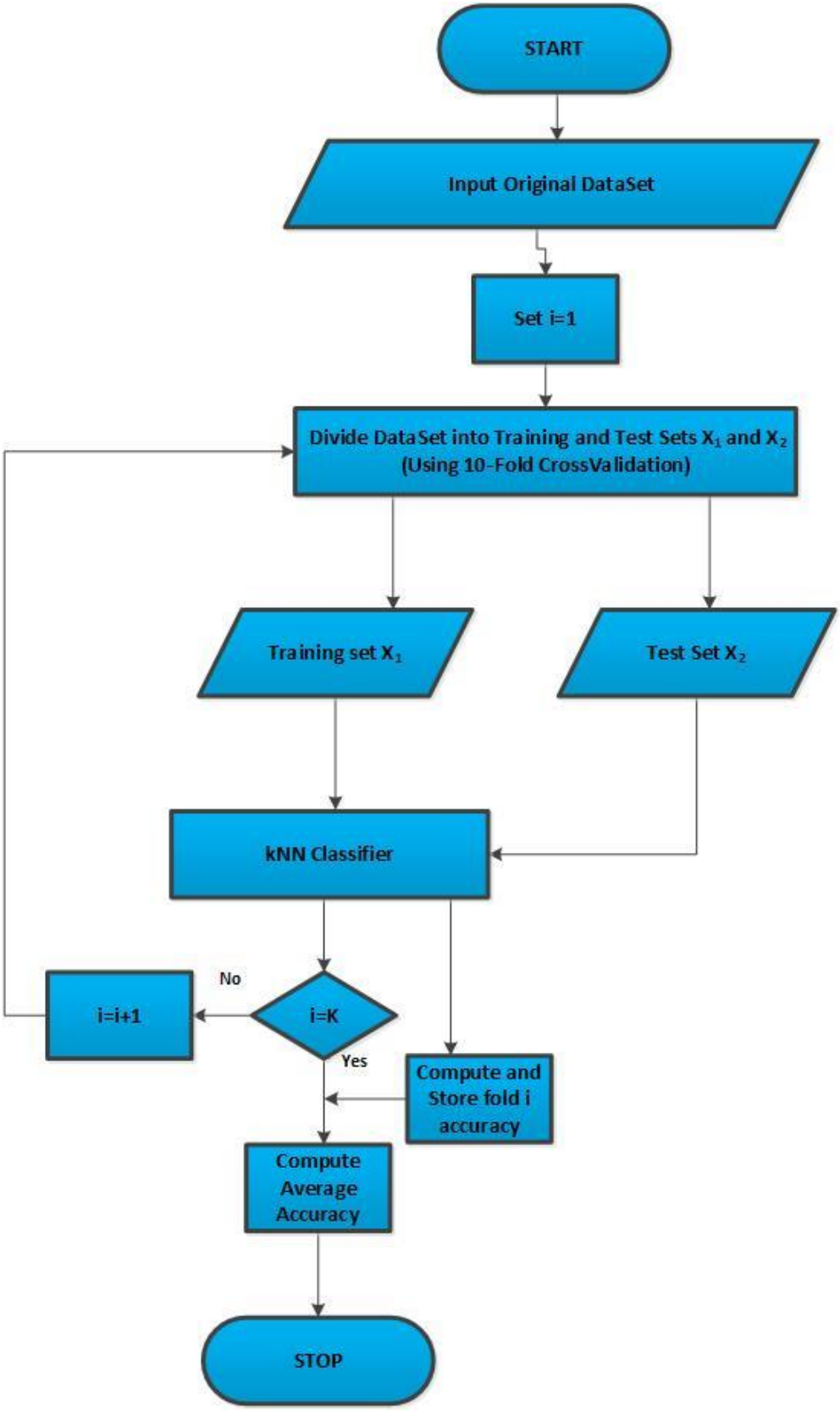

Figure 11. 10-fold cross validation partition based on kNN classifier (adapated from Babatunde, Armstrong, Leng & Diepeveen, 2014c)

Table 3. Experimental results

S/N	Classification model	Accuracy
1	kNN + Original feature set	84.98%
2	kNN + GA_PCA-based features	88.75%

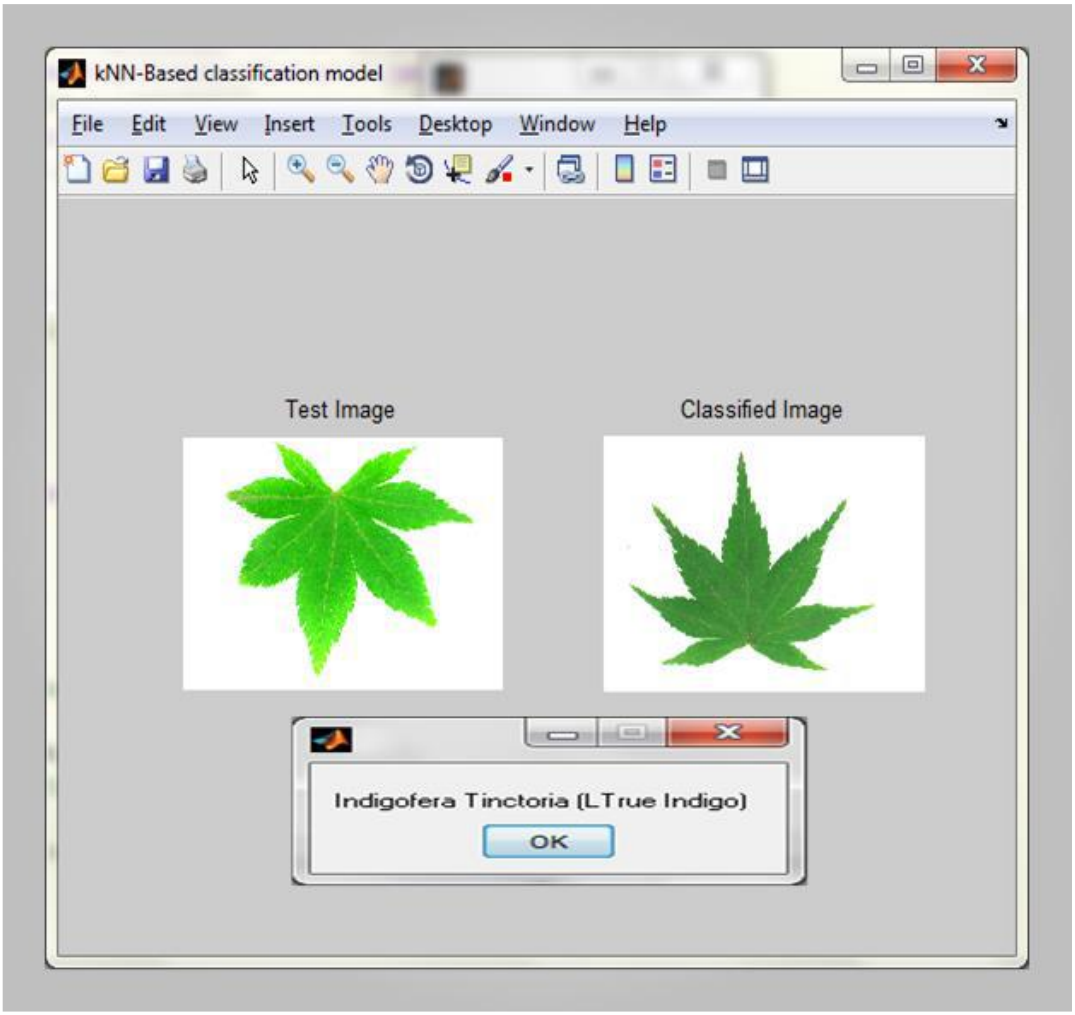

Figure 11. kNN-based classification model

6. Conclusion

Precision agriculture is a multidisciplinary field, integrating various disciplines such as agronomy, computer science, statistics, economics, environmental science, automatic control, telecommunications and microelectronics. The purpose of this paper is to emphasize the role of image processing methods in precision agriculture. This work centers on the application of kNN and genetically selected principal components (PCs) for the development of computer-based vision systems for the identification of plant species. The major new idea involved in this work is the application of GA to automatically select the minimum PC needed to achieve improved accuracy. The original feature space was a 1907 x 112 matrix of real numbers (PCs) while the genetically selected PCs were a 1907 x 41 matrix of real numbers (PCs). The approach in this work can be embedded in agricultural robots in distinguishing weeds from crops. This work can be used in automatic identification of farm animals and other applications involving image processing techniques.

References

Abdul, K., Lukito, E.N., Adhi, S. & Santosa, P.I (2012) Experiments of zernike moments for leaf identification. Journal of Theoretical and Applied Information Technology, 41(1), pp. 83-93.

Andreas, B., Asuka, K., Marion, B., Nick, M., Guido, S. & Andrew, F. (2010) Leafprocessor: a new leaf phenotyping tool contour bending energy and shape cluster analysis. NewPhytologist, 187, pp. 251-261. doi: 10.1111/j.1469-8137.2010.03266.x

Arora, A., Gupta, A., Bagmar, N., Mishra, S. & Bhattacharya, A. (2012) A plant identication system using shape and morphological features on segmented leaves: Teamiitk,clef 2012. Department of Computer Science and Engineering ,Indian Institute of Technology, Kanpur, India and Department of Computer Science and Engineering, University of Florida,Gainesville, USA, pp. 1-14.

Babatunde, O., Armstrong, L., Jinsong, L. & Dean, D. (2014a) On the application of genetic probabilistic neural networks and cellular neural networks in precision agriculture. Asian Journal of Computer and Information Systems, 2(4), pp. 90-100.

Babatunde, O., Armstrong, L., Leng, J. & Diepeveen, D. (2014b) A genetic algorithm-based feature selection. International Journal of Electronics Communication and Computer Engineering, 5, pp. 889–905.

Babatunde, O., Armstrong, L., Leng, J. & Diepeveen, D. (2014c) Zernike moments and genetic algorithm: Tutorial and application. British Journal of Mathematics and Computer Science, 4(15), pp. 2217-2236. doi: 10.9734/bjmcs/2014/10931

Babatunde, O., Armstrong, L., Leng, J. & Diepeveen, D. (2015) A survey of computer-based vision systems for automatic identification of plant species. Journal of Agricultural Informatics, 6(1), pp. 61-71. doi: 10.17700/jai.2015.6.1.152

Jyotismita, C. & Ranjan, P. (2011) Plant leaf recognition using shape based features and neural network classifiers. International Journal of Advanced Computer Science and Applications (IJACSA), pp. 41-47. doi: 10.14569/ijacsa.2011.021007

MathWorks. (2013) Genetic algorithm. Global Optimization Toobox.

Meeta, K., Mrunali,K.,Shubhada,P.,Prajakta,P., & Neha, B. (2012) Survey on techniques for plant leaf classification. International Journal of Modern Engineering Research (IJMER),1(2), pp. 538-544.

McBratney, A., Whelan, B., Ancev, T. (2005) Future Directions of Precision Agriculture. Precision Agriculture, 6, pp. 7-23. doi: 10.1007/s11119-005-0681-8

Nixon, M.S. & Aguado, A.S. (2002) Feature extraction and image processing. Academic Press.

Panagiotis, T. (2005). Plant leaves classification based on morphological features and a fuzzy surface selection techniques. Fifth International Conference on Pattern Recognition, Greece, pp. 365-370.

Sivanandam, S. N. & Deepa, S. N. (2008) Introduction to genetic algorithms. Springer-Verlag , Berlin, Heidelberg. doi: 10.1007/978-3-540-73190-0

Valliammal, N. and Geethalakshmi, S. N. (2011b). Hybrid image segmentation algorithm for leaf recognition and characterization. International Conference on Process Automation, Control and Computing (PACC), pp. 1-6. doi: 10.1109/pacc.2011.5978883

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. In Signal Processing and Information Technology, 2007 IEEE International Symposium on (pp. 11-16). IEEE doi: 10.1109/isspit.2007.4458016

Zalikha, Z., Puteh, S., Itaza, S, & Mohtar, A. (2011) Plant identification using moment invariants and general regression neural network. 11th International Conference on Hybrid Intelligent Systems (HIS), pp. 430-435. doi: 10.1109/his.2011.6122144