International Journal of Computer Applications Technology and Research
Volume 2– Issue 5, 557 - 586, 2013, ISSN: 2319–8656

Analysis of Morphology Based Horticultural Features through Clustering Methods

K.Deb
Department of Computer Science & Engineering
Jadavpur University
Kolkata,India

A.Hazra
Department of Computer Science & Engineering
Jadavpur University
Kolkata,India

S.Kundu
Department of Computer Science & Engineering
Jadavpur University
Kolkata,India

P.Hazra
Faculty of Horticulture, Bidhan Chandra Krishi Viswavidyalaya,
Kalyani,Nadia,India

Abstract: Cluster analysis is a prime Pattern Recognition method used to categorize sample patterns in a population by means of forming different clusters by assigning cluster memberships to the sample patterns depending on the feature similarity relationship among different patterns. Patterns displaying dissimilar feature values are assigned different cluster memberships whereas patterns carrying similar feature values are placed into same cluster. Searching the relationship among horticultural data has become a major research area in Pattern Recognition. In this paper we have used the morphological features for describing the characteristics of Tomato leaves and fruits belonging to different classes. Morphological feature values are extracted from different tomato leaf and fruiting habit samples to analyze through K-Means and Two-step clustering techniques to segment leaf and fruit samples into separate clusters according to their species owing to categorize them. Our experimentation also compares and discusses about the importance of the features which are obtained through K-Means and Two-step Clustering technique, may be useful for leaf and fruit species categorization.

Keywords: Cluster analysis, K-Means Clustering, Two-step clustering, Pattern Recognition, Horticulture, Morphological Feature.

1. INTRODUCTION

Economic growth of a nation depends highly on its agricultural and horticultural development. As different cultivars may need specifically different cultivation processes for better growth and quality development hence it is very much needed to identify the horticultural cultivars belonging to different classes independently, so that appropriate cultivation means can be applied for specific cultivar. Thus horticultural species categorization is a very important task for cultivation. This task of categorization is easy to perform for the crops with small number of species variations, but the task becomes a tough one if huge numbers of crop species are to be dealt with. Horticultural vegetable like tomato have large variety of species found all around the globe. The high variation of morphological feature values of tomato leaf and fruit among different tomato species is a prominent indicator of species diversity upon which cluster analysis would be applied to form different tomato-species clusters and thus categorizing the tomato species depending on their cluster memberships. This automatic method of species categorization through clustering exhibits high level of accuracy and requires trifle time compare to manual process. Cluster analysis also produces some distinguishing results through which the feature importance of leaves and fruits could be predicted which may be of very useful while classifying the particular plant species.

2. BASIC CLUSTER ANALYSIS

Cluster analysis is an important analytical procedure used for the purpose of analysing data. Cluster analysis is widely used in different research areas like machine learning, pattern recognition, market research, digital image processing, Biology etc. Basically, clustering divides the sample entities into different groups called clusters depending on similarity present between entities. Entities placed into same cluster, bear great deal of similar features where as entities belonging to different clusters don’t have that much of feature similarity like same cluster entities. All the member entities of a cluster can be represented by the cluster centre of that cluster. Now with the compact cluster formation, obtaining information from the original entity set can be sufficiently reduced to collecting information about a small number of clusters. Information obtained from the clusters can be very effective for purposes like entity classification, identification etc.

Clustering algorithms can be classified into categories such as: Partitional Clustering and Hierarchical Agglomerative Clustering. We discuss these two clustering techniques briefly in the following-

Partitional Clustering starts with some initial clusters. For each of the initial clusters, a cluster center is calculated by fulfilling the optimality condition. Sample objects are placed in different clusters depending on the smallest distance criterion i.e. a sample object is placed in that cluster whose cluster center is minimum distance away from the sample object. Sample input data are partitioned into the initial clusters. In the next step, cluster centers are recalculated and objects are again placed in different clusters depending on the new calculated cluster centers. This process of cluster center recalculation and placing the objects in clusters continues unless the placement of the objects in the clusters remains unaltered between two successive rotations.

Hierarchical Agglomerative Clustering algorithm starts with some single clusters depending on the size of the input data set. Number of initial single clusters is equal to the size of the input data set and each of the input patterns belongs to different cluster. Now as the algorithm moves, at each of the successive steps, merging of the cluster pairs having highest level of attribute similarities is performed.

3. APPLIED CLUSTERING METHODS

3.1 K-means Cluster Analysis

K-Means cluster analysis falls into the category of partional clustering algorithm. K-Means cluster analysis is used for analysing the feature data set.

Let's assume that P number of sample patterns is to be clustered. A pattern set $D=\{d_1,d_2,..,d_P\}$represents the sample patterns .The characteristics of each sample pattern is represented by Z number of features, which constitute the feature set $F=\{f_1,f_2,\ldots\ldots,f_Z\}$. Now for each of the Z features, P different feature values are obtained from each of the P sample patterns. The feature values associated to a feature f_X forms the individual feature value set $IFV_X=\{ifv_{1X},ifv_{2X},...,ifv_{NX},...,ifv_{PX}\}$,of size P, where ' ifv_{NX}', an element of the set IFV_X, denotes the feature value of N th sample pattern with respect to feature f_X. K-Means cluster analysis is done on each IFV, to place P patterns in k(user given value) different clusters depending on the values of the elements of the IFV. K-Means clustering initially selects k patterns out of P patterns as initial clusters. Each cluster is represented by a cluster center. The value of each initial cluster center will be one of the elements of the IFV chosen randomly with uniqueness condition that same IFV element can't be placed into more than one initial clusters. Also the cluster membership of a particular IFV element remains same till the end of the clustering process.

Let's consider that K-Means clustering is applied to IFV_X. This will lead to the formation of k clusters each having a cluster center. Let's consider that the cluster center of 'i' th cluster is denoted by CC_i. Let's denote the value of cluster center CC_i by VCC_i. Now 'M' th data pattern d_M will be placed into the 'i' th cluster by satisfying the condition $Dis(d_M ,CC_i) < Dis(d_M,CC_j)$,for all $j\neq i$, where $Dis(d_M,CC_i)$ is the distance between the data pattern d_M and the 'i' th cluster center CC_i and $Dis(d_M,CC_j)$ is the distance between d_M and another 'j' th cluster center CC_j. Now, $Dis(d_M,CC_i)$ can be calculated as per the following equation-

$$Dis(d_M,CC_i) = | ifv_{MX} - VCC_i | \quad (1)$$

The values of k cluster centers will be recalculated again and again unless no new member is placed in clusters. The updated value of a cluster center is the calculated average of the values member elements of the cluster. So VCC_i is updated as -

$$VCC_i = (vme_{1i} + vme_{2i} + \ldots + vme_{li} + \ldots + vme_{si}).(1/s) \quad (2)$$

, where 'vme_{1i}' denotes the value of 'l' th member element of 'i' th cluster having 's' number of member elements of the 'i' th cluster. Basically value of each member element is an element of set IFV_X.

In the following, we summarize different steps of K-Means algorithm done on the set IFV_X -

1) Randomly choose k number of initial cluster centers out of P elements of the IFV_X set.
2) Place each pattern from pattern set D in the cluster whose cluster center is closest to it by calculating the pattern-cluster distance as per equation (1).
3) Recompute cluster centers as per equation (2), depending on the recent placement of the elements into the cluster and reassign the elements to its closest cluster based on the newly computed centers.
4) Repeat step 2 and 3 until there is no alteration in the cluster memberships.

3.2 Two step Cluster Analysis

Two-step cluster analysis belongs to the class of Hierarchical Agglomerative Clustering. Consider the pattern set D of size P and feature set F of size Z as mentioned in section 3.1. Values of Z features are extracted from each pattern. Hence values related to Z features, extracted from I th pattern d_I, forms the values of features set $VF_I=\{vf_{I1},vf_{I2},\ldots,vf_{IJ},\ldots vf_{IZ}\}$,where vf_{IJ} is the value of J th feature extracted from I th pattern. Each pattern is represented by its VF set. Now the values of all features extracted from all P patterns build a set of all values of features $AVF=\{VF_1,VF_2,\ldots VF_I\ldots,VF_P\}$. Two-step cluster analysis is performed on the set AVF .Two-step clustering initially forms some sub-clusters and places P patterns into them. Each pattern is described by its VF values, so each sub-cluster contains the VF value of the member pattern. Let's consider that sc_A and sc_B are two sub-clusters containing pattern d_I and d_K respectively. Now the distance between sc_A and sc_B is calculated by calculating the Euclidean Distance between VF_I and VF_K, denoted by $ED(VF_I,VF_K)$,in the following equation-

$$ED(VF_I,VF_K)=((vf_{I1} - vf_{K1})^2+(vf_{I2} - vf_{K2})^2+..+(vf_{IZ} - vf_{KZ})^2)^{1/2} \quad (3)$$

The Two-step algorithm operates on AVF set in the following manner-

1) Place all the P sample patterns into different sub clusters depending on the values of the features set (VF) of each pattern. So each sub-cluster contain the VF set of a pattern.
2) Calculate the distance between sub clusters using equation (3). Merge two nearest sub-clusters(clusters with minimum distance between each other) into one cluster to form new clusters.

Repeat the process of nearest cluster merging between the new clusters until desired number of clusters are formed.

4. EXPERIMENTAL RESULTS

In this paper we have used the best selected morphological features [1] to perform K-Means and Two-step clustering with the help of IBM SPSS statistics 20 data mining tool and there by comparing the cluster building abilities of these features. This comparison will give better visibility about the impact of the features in machine vision solutions. The morphological features used in our experiment are listed below -

Leaf Features

1) Major Axis
2) Minor Axis
3) Aspect Ratio
4) Eccentricity
5) Area
6) Rectangularity

7) Diameter
8) Compactness
9) Perimeter Ratio of Major Axis-Minor Axis
10) Perimeter Ratio of Diameter
11) Concavity
12) R-Factor

Fruit Features

1) Branch Length
2) Branch Width
3) Length Width Ratio
4) Area
5) Perimeter
6) Equivalent Diameter
7) Rectangularity
8) Diameter
9) Perimeter Ratio of Branch Length-Branch Width
10) Perimeter Ratio of Diameter
11) Convexity
12) Solidity
13) On Pixels
14) Narrow-Factor

4.1 K-Means Clustering Results

K-Means clustering is performed individually on feature values, related to each individual leaf and fruit features, extracted from all sample leaf and fruit patterns.

K-Means process starts by randomly selecting 15 and 14 initial cluster centers of individual leaf features (Table 1) and individual fruit features (Table 2) respectively.

Table 1- Randomly chosen Initial cluster centers for individual leaf features

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Major Axis	479	390	419	339	370	435	587	503

Initial Cluster Centers

	Cluster						
	9	10	11	12	13	14	15
Major Axis	599	533	519	575	621	654	557

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Minor Axis	167	321	216	197	254	182	279	291

Initial Cluster Centers

	Cluster						
	9	10	11	12	13	14	15
Minor Axis	150	303	366	404	266	230	245

Initial Cluster Centers

	Cluster						
	1	2	3	4	5	6	7
Aspect Ratio	2.523810	2.795276	3.086207	3.264000	3.175000	2.081340	3.462264

Initial Cluster Centers

	Cluster						
	8	9	10	11	12	13	14
Aspect Ratio	1.912371	2.939850	4.360000	1.602459	1.423792	2.642336	2.410072

Initial Cluster Centers

	Cluster
	15
Aspect Ratio	2.260116

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Eccentricity	.933819	.837966	.943856	.803081	.924178	.874209	.957381	.852387

Initial Cluster Centers

	Cluster						
	9	10	11	12	13	14	15
Eccentricity	.885553	.824524	.781392	.711832	.893455	.914260	.973342

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Area	279909.125	321362.000	315217.375	337626.625	343288.250	330031.625

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Area	298127.250	301354.125	271183.500	276812.625	290063.875	262710.500

Initial Cluster Centers

	Cluster		
	13	14	15
Area	285034.125	304995.625	295351.000

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Rectangularity	.525200	.252211	.287069	.202822	.320442	.369372	.551564	.426617

Initial Cluster Centers

	Cluster						
	9	10	11	12	13	14	15
Rectangularity	.621876	.586994	.792047	.883636	.504833	.463316	.403556

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Diameter	414	316	398	441	356	492	470	614

Initial Cluster Centers

	Cluster						
	9	10	11	12	13	14	15
Diameter	597	650	584	564	545	515	530

Initial Cluster Centers

	Cluster						
	1	2	3	4	5	6	7
Compactness	1.530433	2.035084	1.588200	1.205719	1.857097	1.249416	1.063003

Initial Cluster Centers

	Cluster						
	8	9	10	11	12	13	14
Compactness	1.350639	.982277	1.284001	1.439500	1.319113	1.104743	1.375947

Initial Cluster Centers

	Cluster
	15
Compactness	1.157990

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Perimeter Ratio of Major Axis-Minor Axis	6.694405	7.195364	7.080442	5.983089	7.969643	7.500000

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Perimeter Ratio of Major Axis-Minor Axis	5.771539	5.629520	5.404010	6.914864	5.087829	4.831944

Initial Cluster Centers

	Cluster		
	13	14	15
Perimeter Ratio of Major Axis-Minor Axis	6.293952	6.139844	6.528949

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Perimeter Ratio of Diameter	8.718986	13.753165	11.278894	13.200000	12.536517	8.997152

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Perimeter Ratio of Diameter	7.576137	7.902203	8.465625	9.509748	8.228941	10.629945

Initial Cluster Centers

	Cluster		
	13	14	15
Perimeter Ratio of Diameter	10.257154	9.872329	6.920409

Initial Cluster Centers

	Cluster						
	1	2	3	4	5	6	7
Concavity	91802.875	50350.000	56494.625	34085.375	28423.750	41680.375	73584.750

Initial Cluster Centers

	Cluster					
	8	9	10	11	12	13
Concavity	70357.875	100528.500	94899.375	81648.125	108918.500	86677.875

Initial Cluster Centers

	Cluster	
	14	15
Concavity	66716.375	76361.000

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
R-factor	897.855072	1176.303797	933.949749	1143.729231	1044.134831	745.879518

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
R-factor	791.156550	571.898383	622.190955	769.043478	658.595745	842.285714

Initial Cluster Centers

	Cluster		
	13	14	15
R-factor	682.181818	726.193548	703.500000

Table 2- Randomly chosen Initial cluster centers for individual fruit features

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Branch Length	508.861407	631.948827	484.844350	654.464819	574.908316	691.991471

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Branch Length	447.317697	598.925373	586.916844	537.381663	553.893390	615.437100

Initial Cluster Centers

	Cluster	
	13	14
Branch Length	520.869936	673.978678

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Branch Width	411.292111	358.754797	475.837953	321.228145	382.771855	295.710021

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Branch Width	447.317697	496.852878	460.827292	256.682303	424.801706	340.742004

Initial Cluster Centers

	Cluster	
	13	14
Branch Width	235.667377	306.217484

Initial Cluster Centers

	Cluster						
	1	2	3	4	5	6	7
Length width Ratio	2.491228	1.784404	1.017065	2.039409	1.491857	2.192308	1.206790

Initial Cluster Centers

	Cluster						
	8	9	10	11	12	13	14
Length width Ratio	1.414894	2.821656	2.319797	1.306569	1.650655	1.577061	1.926471

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Area	175316	207688	186587	240592	292808	269369	200254	227813

Initial Cluster Centers

	Cluster					
	9	10	11	12	13	14
Area	248800	134572	256387	193219	160739	213099

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Perimeter	4396	4808	4946	5793	5720	6427	5205	4603

Initial Cluster Centers

	Cluster					
	9	10	11	12	13	14
Perimeter	5056	5475	3815	5319	4530	5569

Initial Cluster Centers

	Cluster				
	1	2	3	4	5
Equivalent Diameter	472.461197	514.234433	487.411966	553.472282	610.585559

Initial Cluster Centers

	Cluster				
	6	7	8	9	10
Equivalent Diameter	585.637892	504.946683	538.572670	562.834324	413.934674

Initial Cluster Centers

	Cluster			
	11	12	13	14
Equivalent Diameter	571.350950	495.998216	452.393334	520.889542

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Rectangularity	1.51555	.98869	1.35499	.56056	.74742	.89254	1.25091	1.46348

Initial Cluster Centers

	Cluster					
	9	10	11	12	13	14
Rectangularity	1.06472	1.12193	1.58324	1.20264	2.47732	1.41198

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Diameter	421.799574	483.343284	522.371002	570.405117	639.454158	675.479744

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Diameter	597.424307	619.940299	580.912580	661.970149	505.859275	697.995736

Initial Cluster Centers

	Cluster	
	13	14
Diameter	609.432836	447.317697

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Perimeter Ratio of Branch Length-Branch Width	4.430537	4.718868	5.148424	4.235480	5.946467	5.038155

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Perimeter Ratio of Branch Length-Branch Width	4.880447	3.996112	5.745399	5.286098	6.160609	3.301852

Initial Cluster Centers

	Cluster	
	13	14
Perimeter Ratio of Branch Length-Branch Width	5.415092	5.539335

Initial Cluster Centers

	Cluster					
	1	2	3	4	5	6
Perimeter Ratio of Diameter	10.422011	9.085497	8.155919	8.447897	9.686931	7.963229

Initial Cluster Centers

	Cluster					
	7	8	9	10	11	12
Perimeter Ratio of Diameter	9.372221	5.772676	8.703547	7.397360	8.894112	6.618560

Initial Cluster Centers

	Cluster	
	13	14
Perimeter Ratio of Diameter	7.707393	11.890878

Initial Cluster Centers

	Cluster						
	1	2	3	4	5	6	7
Convexity	84.55687	79.52760	77.31115	72.58582	64.13240	57.83600	70.96449

Initial Cluster Centers

	Cluster						
	8	9	10	11	12	13	14
Convexity	82.05563	80.74126	68.10407	66.74663	69.29754	87.08688	74.88155

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
Solidity	.471645	.558735	.501968	.647254	.787728	.724672	.538734	.612875

Initial Cluster Centers

	Cluster					
	9	10	11	12	13	14
Solidity	.669336	.400704	.689746	.519809	.434213	.573290

Initial Cluster Centers

	Cluster							
	1	2	3	4	5	6	7	8
On-Pixels	175227	207604	186498	240498	292731	269295	200173	227711

Initial Cluster Centers

	Cluster					
	9	10	11	12	13	14
On-Pixels	248725	134514	256313	193144	160675	212979

Initial Cluster Centers

	Cluster						
	1	2	3	4	5	6	7
Narrow-Factor	.891534	.819477	1.250774	.927954	.953964	.986842	1.000000

Initial Cluster Centers

	Cluster						
	8	9	10	11	12	13	14
Narrow-Factor	1.146341	1.015284	1.065491	1.038560	.974359	1.094017	1.111959

Table 3 and Table 4 represent the iteration history of the K-Means clustering on individual leaf and fruit features respectively. Iteration history shows the number of times the clustering iterates before completion. Clustering algorithm completes when there is a small change or no change in the cluster centers, there by achieving the convergence.

Table 3- Iteration History of clustering on 12 individual leaf features

Table 3.1- Iteration History of clustering on Major Axis

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.000	2.150	.000	.000	.000	.530	2.502
2	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.000	3.002	.000	4.203	4.003	.000	4.000E-007
2	.000	1.126	.000	.000	.000	.000	1.334
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 12.009.

Table 3.2- Iteration History of clustering on Minor Axis

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.375	.000	.790	.461	1.501	.375	1.001	.000
2	2.377	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	3.753	.000	.000	.000	3.002	1.876	3.002
2	1.751	.000	.000	.000	.000	1.876	1.501
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 9.006.

Table 3.3- Iteration History of clustering on Aspect Ratio

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.021	.000	.015	.000	.016	.028	.013	.041
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.000	.000	.053	.000	.023	.005	.001
2	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .089.

Table 3.4- Iteration History of clustering on Eccentricity

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.002	.001	.002	.001	.002	.002	.002	.004
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.002	.000	.000	.000	.003	2.325E-005	.000
2	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .008.

Table 3.5- Iteration History of clustering on Area

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	1280.688	804.875	561.531	.000	.000	1761.583	92.188	605.896
2	.000	.000	.000	.000	.000	.000	473.604	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	109.938	.000	477.083	.000	.000	49.875	788.925
2	.000	.000	.000	.000	.000	.000	513.044
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 2776.250.

Table 3.6- Iteration History of clustering on Rectangularity

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.007	.002	.004	.001	.003	.001	.000	.006
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.004	.000	.000	.000	.006	.000	.007
2	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .020.

Table 3.7- Iteration History of clustering on Diameter

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	4.500	.000	4.503	.000	1.629	6.004	3.002
2	.000	.000	.000	.000	.000	1.329	1.501	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.000	.000	4.878	1.001	.429	.500	1.876
2	.000	.000	1.126	4.003	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 13.510.

Table 3.8- Iteration History of clustering on Compactness

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.009	.000	.001	.000	8.000E-006	.005	.000
2	.000	.000	.000	.000	.000	.000	.008	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.017	8.050E-005	.000	.004	.001	.000	.002
2	.000	.000	.000	.000	.007	.000	.000
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is .025.

Table 3.9- Iteration History of clustering on Perimeter Ratio of Major Axis-Minor Axis

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.034	.000	.000	.018	.020	.025	.039	.007
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.044	.000	.023	.000	.012	.027	.058
2	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .115.

Table 3.10- Iteration History of clustering on Perimeter Ratio of Diameter

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.014	.000	.004	.000	.000	.080	.072	.013
2	.040	.000	.000	.000	.000	.026	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	.034	.008	.009	.000	.048	.028	.086
2	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is .237.

Table 3.11- Iteration History of clustering on Concavity

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	1280.688	804.875	561.531	.000	.000	1761.583	92.188	605.896
2	.000	.000	.000	.000	.000	.000	473.604	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	109.938	.000	477.083	.000	.000	49.875	788.925
2	.000	.000	.000	.000	.000	.000	513.044
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 2776.250.

Table 3.12- Iteration History of clustering on R-Factor

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.000	.000	.000	.000	1.634	3.755	.000
2	.000	.000	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers						
	9	10	11	12	13	14	15
1	3.273	1.367	7.984	8.423	.602	3.522	4.450
2	5.889	.000	2.004	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 19.686.

Table 4- Iteration History of clustering on 14 individual fruit features

Table 4.1- Iteration History of clustering on Branch-Length

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	1.501	.000	5.004	3.753	4.503	.000	1.501
2	.000	.000	.000	1.376	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.563	.000	2.502	3.002	3.002	3.002
2	.000	.000	.000	.000	.000	1.501
3	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 12.009.

Table 4.2- Iteration History of clustering on Branch-Width

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	5.254	4.003	3.333E-007	4.503	3.377	.000	3.603	6.004
2	.000	.000	3.753	.000	.000	.000	3.603	1.876
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	1.501	.000	.751	3.377	.000	1.501
2	1.501	.000	1.951	.000	.000	.000
3	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 10.507.

Table 4.3- Iteration History of clustering on Length-Width Ratio

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.021	.013	.004	.010	.000	.025	.023
2	.000	.000	.000	.000	.000	.000	.007	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.000	.000	.019	.004	.016	.021
2	.000	.000	.009	.000	.000	.000
3	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is .074.

Table 4.4- Iteration History of clustering on Area

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	649.813	420.825	1662.792	1128.833	.000	2781.417	.000	2060.188
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	529.025	.000	1237.800	.000	331.438	2533.313
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 5410.500.

Table 4.5- Iteration History of clustering on Perimeter

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	31.000	9.000	.000	.000	.000	2.889	37.500
2	.000	.000	.000	.000	.000	.000	8.486	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	24.333	4.667	20.500	10.444	15.000	7.500
2	14.917	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 73.000.

Table 4.6- Iteration History of clustering on Equivalent-Diameter

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.877	.516	2.165	1.309	.000	3.038	.000	2.422
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.604	.000	1.366	.000	.466	3.078
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 6.655.

Table 4.7- Iteration History of clustering on Rectangularity

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.003	.016	.014	.013	.022	.002	.017	.000
2	.000	.006	.000	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.004	.012	.006	.004	.000	.001
2	.012	.000	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is .048.

Table 4.8- Iteration History of clustering on Diameter

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.000	2.252	2.502	5.000E-007	1.876	3.002	.000
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.375	.188	.000	3.002	1.501	.000
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 10.507.

Table 4.9- Iteration History of clustering on Perimeter Ratio of Branch Length-Branch Width

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.039	.010	.000	.009	.000	.023	.000
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.007	.024	.024	.000	.013	.000
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .110.

Table 4.10- Iteration History of clustering on Perimeter Ratio of Diameter

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.014	.045	.049	.000	.049	.037	.000
2	.000	.000	.000	.000	.000	.016	.000	.000
3	.000	.000	.000	.000	.000	.015	.000	.000
4	.000	.000	.000	.000	.000	.012	.000	.000
5	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.016	.014	.015	.000	.009	.000
2	.000	.000	.000	.000	.015	.000
3	.000	.000	.000	.000	.016	.000
4	.000	.000	.000	.000	.021	.000
5	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 5. The minimum distance between initial centers is .191.

Table 4.11- Iteration History of clustering on Convexity

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.529	.029	.505	.370	.426	.000	.096	.270
2	.000	.000	.000	.000	.000	.000	.199	.000
3	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.055	.268	.090	.302	.000	.136
2	.000	.000	.000	.148	.000	.000
3	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 1.193.

Table 4.12- Iteration History of clustering on Solidity

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.002	.001	.004	.003	.000	.007	.000	.006
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.001	.000	.003	.000	.000	.007
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .015.

Table 4.13- Iteration History of clustering on On Pixels

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	649.500	418.800	1671.667	1112.667	.000	2790.667	.000	2075.250
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	530.200	.000	1240.000	.000	316.000	2549.500
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 5375.000.

Table 4.14- Iteration History of clustering on Narrow-Factor

Iteration History[a]

Iteration	Change in Cluster Centers							
	1	2	3	4	5	6	7	8
1	.000	.005	.000	.001	.000	.000	.001	.005
2	.000	.000	.000	.000	.000	.000	.000	.000

Iteration History[a]

Iteration	Change in Cluster Centers					
	9	10	11	12	13	14
1	.001	.000	.005	.000	.006	.005
2	.000	.000	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is .012.

Information regarding the membership of 15 leaf and 14 fruit clusters build through K-Means Clustering is shown by Table-5 and Table 6. The case number field signifies the sample leaf/fruit pattern number. The field called cluster is the cluster number in which a pattern is placed and distance field gives the distance between the pattern and the cluster center, in which the pattern is placed.

Table 5-Cluster membership of leaf clusters build from 12 individual leaf features

Table 5.1-Cluster membership for Major Axis

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	2	.000
3	3	2.150
4	4	.000
5	5	.000
6	6	.000
7	7	1.148
8	15	1.334
9	9	.000
10	10	1.126
11	7	.530
12	12	4.203
13	15	1.334
14	7	2.031
15	15	1.668

Cluster Membership		
Case Number	Case Number	Case Number
16	10	.375
17	15	7.339
18	13	1.001
19	15	4.670
20	12	1.201
21	10	7.881
22	8	1.001
23	10	4.128
24	15	1.668
25	15	.167
26	11	.000
27	10	4.878
28	8	2.502
29	7	.530
30	12	4.803

Cluster Membership		
Case Number	Cluster	Distance
31	11	.000
32	13	4.003
33	7	2.472
34	12	4.203
35	7	.530
36	10	5.629
37	10	6.380
38	15	3.169
39	14	.000
40	3	2.150
41	12	4.803
42	10	8.631
43	8	3.502
44	15	1.334
45	13	5.004

Table 5.2-Cluster membership for Minor Axis

Cluster Membership		
Case Number	Cluster	Distance
1	3	2.637
2	3	1.363
3	3	.363
4	4	4.900
5	4	7.100
6	3	6.363
7	7	1.001
8	8	.000
9	7	.500
10	10	.000
11	11	.000
12	12	.000
13	13	3.002
14	15	3.002
15	15	4.503
16	1	5.504
17	9	3.502
18	14	2.842E-014
19	6	.375
20	6	1.876
21	9	2.001
22	1	2.001
23	4	6.465
24	3	5.294
25	14	2.842E-014
26	3	.790
27	2	.000
28	3	6.715
29	4	2.542
30	1	3.502
31	5	3.333E-007
32	5	1.501
33	5	1.501
34	14	2.842E-014
35	13	3.002
36	15	1.501
37	4	8.546
38	6	3.377
39	9	5.504
40	4	.461
41	3	.790
42	4	1.962
43	7	.500
44	3	5.294
45	6	5.629

Table 5.3-Cluster membership for Aspect Ratio

Cluster Membership		
Case Number	Cluster	Distance
1	15	.062
2	8	.049
3	8	.077
4	11	.022
5	8	.076
6	6	.028
7	6	.001
8	8	.041
9	6	.048
10	8	.104
11	11	.053
12	12	.000
13	6	.013
14	1	.034
15	15	.029
16	3	.015
17	7	.013
18	13	.028
19	3	.028
20	5	.016
21	7	.008
22	3	.044
23	2	.000
24	1	.035
25	14	.017
26	14	.012
27	11	.031
28	14	.005
29	9	.000
30	7	.022
31	6	.050
32	1	.023
33	15	.066
34	1	.001
35	15	.001
36	15	.033
37	13	.023
38	5	.016
39	10	.000
40	6	.044
41	13	.047
42	13	.042
43	8	.039
44	1	.021
45	4	.000

Table 5.4-Cluster membership for Eccentricity

Cluster Membership		
Case Number	Cluster	Distance
1	13	.006
2	2	.001
3	8	.002
4	4	.001
5	8	.002
6	6	.000
7	9	.004
8	8	.004
9	9	.002
10	10	.000
11	11	.000
12	12	.000
13	6	.002
14	14	2.325E-005
15	13	.003
16	3	.000
17	7	.002
18	5	.002
19	3	.001
20	3	.003
21	7	.001
22	3	.002
23	1	.002
24	14	.005
25	14	.003
26	14	.005
27	4	.001
28	14	.004
29	3	.005
30	7	.001
31	6	.002
32	14	.001
33	13	.006
34	14	.003
35	13	.000
36	13	.003
37	5	.000
38	3	.002
39	15	.000
40	9	.002
41	5	.002
42	1	.002
43	2	.001
44	14	.004
45	7	.004

Table 5.5-Cluster membership for Area

Cluster Membership

Case Number	Cluster	Distance
1	3	482.281
2	2	2642.500
3	3	1969.281
4	4	.000
5	5	.000
6	6	914.792
7	7	381.417
8	8	605.896
9	9	109.938
10	10	.000
11	9	109.938
12	12	.000
13	15	960.406
14	14	952.000
15	15	1301.969
16	3	1168.906
17	2	194.125
18	15	1167.531
19	14	1211.875
20	14	141.375
21	3	1260.469
22	8	283.021
23	3	783.719
24	11	948.458
25	7	565.792
26	14	451.125
27	1	1280.688
28	6	846.792
29	3	1443.969
30	3	693.844
31	8	1026.271
32	7	947.208
33	15	825.969
34	1	1280.688
35	13	.000
36	11	477.083
37	8	710.604
38	2	804.875
39	2	2031.750
40	6	1761.583
41	8	1149.354
42	11	1425.542
43	14	49.875
44	3	561.531
45	8	55.229

Table 5.6-Cluster membership for Rectangularity

Cluster Membership

Case Number	Cluster	Distance
1	5	.008
2	2	.002
3	3	.004
4	4	.001
5	4	.001
6	3	.005
7	7	.000
8	1	.005
9	9	.004
10	10	.000
11	11	.000
12	12	.000
13	13	.006
14	14	.006
15	14	.001
16	3	.015
17	3	.011
18	13	.014
19	5	.011
20	5	.014
21	3	.012
22	3	.004
23	5	.002
24	8	.006
25	8	.002
26	6	.001
27	9	.004
28	5	.003
29	6	.001
30	3	.013
31	8	.003
32	1	.007
33	13	.008
34	14	.005
35	1	.002
36	8	.005
37	6	.001
38	5	.011
39	5	.016
40	2	.002
41	15	.007
42	6	.015
43	14	.000
44	15	.007
45	6	.012

Table 5.7-Cluster membership for Diameter

Cluster Membership

Case Number	Cluster	Distance
1	1	.000
2	2	4.500
3	3	.000
4	2	4.500
5	5	.000
6	6	.308
7	11	1.501
8	8	3.002
9	9	.000
10	7	6.004
11	11	6.004
12	11	4.503
13	13	.429
14	14	2.502
15	15	2.627
16	7	1.501
17	7	3.002
18	8	3.002
19	14	2.001
20	11	3.002
21	15	2.627
22	7	7.505
23	6	1.457
24	4	4.503
25	15	1.876
26	6	3.046
27	13	4.074
28	6	7.549
29	13	2.573
30	6	6.048
31	6	5.960
32	12	3.002
33	12	3.002
34	13	4.074
35	13	5.575
36	15	3.377
37	14	.500
38	13	7.934
39	10	.000
40	4	4.503
41	7	3.002
42	6	2.958
43	6	5.960
44	13	7.934
45	11	2.000E-007

Table 5.8-Cluster membership for Compactness

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	2	.009
3	3	.000
4	2	.009
5	5	.000
6	12	.011
7	7	.013
8	9	.020
9	9	.016
10	6	.021
11	9	.008
12	9	.009
13	13	.010
14	4	.011
15	15	.009
16	12	.007
17	8	.000
18	9	.004
19	4	.004
20	7	.008
21	4	.016
22	12	.004
23	10	.001
24	14	.000
25	15	.006
26	6	8.000E-006
27	13	.006
28	10	8.050E-005
29	15	.016
30	6	.017
31	6	.016
32	13	.010
33	7	.009
34	13	.001
35	13	.007
36	15	.006
37	4	.001
38	15	.002
39	9	.017
40	11	.000
41	10	.009
42	6	.011
43	10	.011
44	15	.009
45	7	.003

Table 5.9-Cluster membership for Perimeter Ratio of Major Axis-Minor Axis

Cluster Membership		
Case Number	Cluster	Distance
1	1	.034
2	2	.000
3	3	.000
4	5	.020
5	5	.020
6	6	.025
7	7	.048
8	8	.027
9	9	.015
10	9	.059
11	11	.023
12	12	.000
13	1	.033
14	14	.018
15	15	.030
16	7	.057
17	14	.014
18	11	.023
19	7	.032
20	9	.044
21	7	.062
22	1	.001
23	14	.043
24	7	.039
25	14	.024
26	13	.050
27	7	.063
28	4	.018
29	13	.065
30	13	.012
31	14	.041
32	13	.046
33	4	.018
34	8	.008
35	14	.027
36	13	.047
37	15	.028
38	8	.018
39	8	.046
40	6	.025
41	15	.058
42	10	.000
43	14	.029
44	13	.004
45	8	.007

Table 5.10-Cluster membership for Perimeter Ratio of Diameter

Cluster Membership		
Case Number	Cluster	Distance
1	3	.004
2	2	.000
3	3	.004
4	4	.000
5	5	.000
6	14	.087
7	1	.142
8	8	.013
9	8	.083
10	10	.059
11	11	.009
12	11	.009
13	14	.161
14	14	.007
15	14	.148
16	1	.054
17	6	.026
18	15	.022
19	9	.007
20	15	.063
21	7	.072
22	10	.008
23	1	.120
24	13	.001
25	6	.024
26	10	.157
27	6	.078
28	9	.034
29	6	.079
30	6	.074
31	10	.073
32	14	.039
33	6	.106
34	9	.027
35	10	.145
36	6	.017
37	10	.128
38	7	.072
39	15	.086
40	13	.048
41	12	.000
42	13	.047
43	14	.028
44	1	.076
45	8	.096

Table 5.11-Cluster membership for Concavity

Cluster Membership		
Case Number	Cluster	Distance
1	3	482.281
2	2	2642.500
3	3	1969.281
4	4	.000
5	5	.000
6	6	914.792
7	7	381.417
8	8	605.896
9	9	109.938
10	10	.000
11	9	109.938
12	12	.000
13	15	960.406
14	14	952.000
15	15	1301.969
16	3	1168.906
17	2	194.125
18	15	1167.531
19	14	1211.875
20	14	141.375
21	3	1260.469
22	8	283.021
23	3	783.719
24	11	948.458
25	7	565.792
26	14	451.125
27	1	1280.688
28	6	846.792
29	3	1443.969
30	3	693.844
31	8	1026.271
32	7	947.208
33	15	825.969
34	1	1280.688
35	13	.000
36	11	477.083
37	8	710.604
38	2	804.875
39	2	2031.750
40	6	1761.583
41	8	1149.354
42	11	1425.542
43	14	49.875
44	3	561.531
45	8	55.229

Table 5.12-Cluster membership for R-Factor

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	2	.000
3	3	.000
4	4	.000
5	5	.000
6	6	3.420
7	11	7.074
8	9	1.591
9	9	9.163
10	10	1.367
11	11	12.022
12	11	.503
13	13	.602
14	14	3.522
15	15	3.454
16	10	8.599
17	7	3.755
18	9	7.571
19	14	2.811
20	11	2.049
21	15	3.454
22	7	3.755
23	6	5.167
24	12	8.423
25	15	2.458
26	6	1.634
27	13	5.083
28	6	8.314
29	13	3.178
30	6	6.101
31	10	8.069
32	11	17.069
33	11	9.988
34	13	5.083
35	13	6.999
36	15	4.450
37	14	.712
38	13	9.871
39	8	.000
40	12	8.423
41	10	6.173
42	6	7.462
43	10	8.069
44	13	9.871
45	11	5.408

Table 6-Cluster membership of fruit clusters build from 14 individual fruit features

Table 6.1-Cluster membership for Branch Length

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	2	1.501
3	3	.000
4	4	6.380
5	4	2.627
6	6	3.002
7	9	.563
8	8	1.501
9	5	2.252
10	8	1.501
11	11	2.502
12	12	3.002
13	5	2.252
14	9	3.565
15	9	5.066
16	11	2.502
17	8	1.501
18	10	.000
19	11	5.004
20	5	3.753
21	13	3.002
22	9	.938
23	14	1.501
24	2	4.503
25	5	3.753
26	8	1.501
27	9	.563
28	14	4.503
29	6	4.503
30	6	.000
31	14	1.501
32	4	.375
33	14	4.503
34	13	3.002
35	9	3.940
36	7	.000
37	2	6.004
38	4	4.128
39	6	1.501
40	9	2.439
41	9	2.439
42	12	3.002

Table 6.2-Cluster membership for Branch Width

Cluster Membership		
Case Number	Cluster	Distance
1	8	5.629
2	2	4.003
3	3	3.753
4	4	4.503
5	5	3.377
6	14	4.503
7	8	2.627
8	8	.375
9	9	1.501
10	7	.300
11	11	1.201
12	9	4.503
13	2	3.502
14	14	1.501
15	2	.500
16	7	4.803
17	7	7.205
18	1	5.254
19	1	5.254
20	12	.375
21	5	7.881
22	3	3.753
23	8	7.881
24	9	.000
25	5	5.629
26	11	.300
27	12	4.128
28	12	.375
29	12	3.377
30	9	3.002
31	7	1.801
32	11	4.803
33	11	.300
34	11	4.203
35	5	5.629
36	7	.300
37	10	.000
38	13	.000
39	6	.000
40	4	1.501
41	4	3.002
42	14	3.002

Table 6.3-Cluster membership for Length Width Ratio

Cluster Membership		
Case Number	Cluster	Distance
1	3	.023
2	2	.002
3	3	.011
4	4	.002
5	2	.031
6	6	.000
7	7	.032
8	7	.015
9	7	.010
10	11	.021
11	11	.030
12	11	.023
13	13	.008
14	14	.021
15	12	.027
16	7	.034
17	11	.002
18	11	.028
19	8	.009
20	12	.004
21	8	.015
22	7	.011
23	11	.016
24	8	.018
25	5	.015
26	8	.023
27	12	.031
28	14	.021
29	4	.005
30	5	.010
31	13	.024
32	13	.016
33	13	.017
34	7	.007
35	5	.005
36	3	.013
37	1	.000
38	9	.000
39	10	.000
40	2	.021
41	2	.012
42	4	.004

Table 6.4-Cluster membership for Area

Cluster Membership		
Case Number	Cluster	Distance
1	4	2646.583
2	8	2053.063
3	2	1081.175
4	9	2388.100
5	8	3303.813
6	6	3369.958
7	4	3775.417
8	8	1675.313
9	9	899.725
10	8	3736.813
11	11	1237.800
12	8	2110.063
13	2	3292.325
14	8	2941.813
15	2	420.825
16	14	2533.313
17	14	2533.313
18	4	1128.833
19	9	3817.650
20	11	3268.700
21	11	4254.175
22	12	.000
23	10	.000
24	3	909.208
25	8	2060.188
26	13	331.438
27	8	420.313
28	1	649.813
29	1	649.813
30	2	932.425
31	3	1662.792
32	3	753.583
33	7	.000
34	13	331.438
35	2	1699.550
36	11	4201.575
37	5	.000
38	6	588.542
39	6	2781.417
40	11	1978.300
41	9	529.025
42	9	1058.850

Table 6.5-Cluster membership for Perimeter

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	8	34.500
3	3	9.000
4	9	25.750
5	5	.000
6	12	49.556
7	7	21.625
8	12	3.444
9	9	39.250
10	10	4.667
11	9	44.750
12	12	34.556
13	9	31.250
14	3	9.000
15	7	5.375
16	12	34.556
17	6	.000
18	14	7.500
19	2	31.000
20	13	15.000
21	13	15.000
22	7	36.375
23	11	20.500
24	8	37.500
25	14	7.500
26	7	3.625
27	12	40.444
28	7	13.625
29	7	20.625
30	10	21.667
31	12	2.444
32	12	39.444
33	7	11.375
34	11	20.500
35	4	.000
36	12	10.444
37	7	6.375
38	10	26.333
39	12	22.444
40	8	30.500
41	8	33.500
42	2	31.000

Table 6.6-Cluster membership for Equivalent Diameter

Cluster Membership		
Case Number	Cluster	Distance
1	4	3.052
2	8	2.413
3	2	1.340
4	9	2.705
5	8	3.882
6	6	3.687
7	4	4.361
8	8	1.967
9	9	1.015
10	8	4.387
11	11	1.366
12	8	2.485
13	2	4.083
14	8	3.465
15	2	.516
16	14	3.078
17	14	3.078
18	4	1.309
19	9	4.311
20	11	3.633
21	11	4.737
22	12	.000
23	10	.000
24	3	1.183
25	8	2.422
26	13	.466
27	8	.487
28	1	.877
29	1	.877
30	2	1.157
31	3	2.165
32	3	.981
33	7	.000
34	13	.466
35	2	2.102
36	11	4.663
37	5	.000
38	6	.649
39	6	3.038
40	11	2.192
41	9	.604
42	9	1.195

Table 6.7-Cluster membership for Rectangularity

Cluster Membership		
Case Number	Cluster	Distance
1	2	.006
2	2	.015
3	10	.007
4	6	.040
5	9	.009
6	5	.043
7	12	.004
8	7	.017
9	9	.016
10	10	.012
11	6	.023
12	12	.004
13	2	.007
14	5	.027
15	2	.024
16	10	.005
17	7	.017
18	6	.024
19	6	.002
20	5	.022
21	5	.008
22	14	.008
23	13	.000
24	1	.003
25	2	.022
26	11	.006
27	6	.004
28	3	.014

Cluster Membership		
Case Number	Cluster	Distance
29	3	.014
30	1	.003
31	11	.006
32	8	.000
33	14	.001
34	14	.009
35	9	.007
36	5	.018
37	4	.013
38	4	.013
39	5	.016
40	5	.022
41	5	.002
42	5	.003

Table 6.8-Cluster membership for Diameter

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	3	2.252
3	13	1.501
4	10	10.695
5	5	6.004
6	6	1.876
7	9	6.380
8	7	3.002
9	9	.375
10	5	4.503
11	9	2.627
12	5	3.002
13	5	10.507
14	10	6.192

Cluster Membership		
Case Number	Cluster	Distance
15	10	2.814
16	5	4.503
17	10	1.313
18	13	1.501
19	3	2.252
20	11	.000
21	2	.000
22	4	2.502
23	10	5.817
24	8	.000
25	4	5.504
26	7	3.002
27	4	8.006
28	6	1.876

Cluster Membership		
Case Number	Cluster	Distance
29	12	3.002
30	12	3.002
31	6	4.878
32	10	.188
33	10	5.817
34	9	4.128
35	5	9.006
36	14	.000
37	5	5.000E-007
38	10	1.313
39	6	8.631
40	13	1.501
41	13	1.501
42	5	10.507

Table 6.9-Cluster membership for Perimeter Ratio of Branch Length-Branch Width

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	2	.039
3	3	.010
4	10	.014
5	13	.040
6	13	.031
7	7	.023
8	7	.009
9	7	.008
10	10	.024
11	10	.010
12	6	.020
13	13	.013
14	14	.000

Cluster Membership		
Case Number	Cluster	Distance
15	13	.031
16	13	.005
17	11	.024
18	5	.067
19	7	.029
20	6	.034
21	6	.000
22	7	.012
23	12	.000
24	4	.000
25	9	.007
26	3	.032
27	9	.083
28	3	.014

Cluster Membership		
Case Number	Cluster	Distance
29	6	.033
30	2	.005
31	2	.032
32	7	.002
33	2	.012
34	8	.000
35	5	.009
36	5	.058
37	9	.076
38	11	.024
39	13	.022
40	6	.021
41	3	.017
42	3	.053

Table 6.10-Cluster membership for Perimeter Ratio of Diameter

Cluster Membership		
Case Number	Cluster	Distance
1	1	.000
2	2	.044
3	3	.045
4	6	.011
5	2	.041
6	6	.091
7	11	.015
8	2	.066
9	9	.016
10	9	.065
11	11	.015
12	4	.049
13	6	.081
14	13	.080

Cluster Membership		
Case Number	Cluster	Distance
15	6	.036
16	4	.049
17	5	.000
18	2	.067
19	2	.014
20	2	.057
21	7	.037
22	2	.010
23	8	.000
24	10	.014
25	5	.000
26	9	.050
27	7	.037
28	13	.095

Cluster Membership		
Case Number	Cluster	Distance
29	13	.080
30	6	.052
31	6	.049
32	6	.119
33	6	.080
34	12	.000
35	2	.117
36	14	.000
37	3	.053
38	3	.098
39	13	.089
40	13	.083
41	13	.060
42	10	.014

Table 6.11-Cluster membership for Convexity

Cluster Membership		
Case Number	Cluster	Distance
1	1	.529
2	2	.029
3	14	.136
4	4	.370
5	5	.426
6	12	.643
7	7	.295
8	12	.045
9	4	.563
10	10	.057
11	4	.639
12	12	.450
13	4	.447
14	14	.136

Cluster Membership		
Case Number	Cluster	Distance
15	7	.073
16	12	.450
17	6	.000
18	11	.090
19	3	.505
20	8	.270
21	8	.270
22	7	.500
23	13	.000
24	9	.055
25	11	.090
26	7	.050
27	12	.515
28	7	.186

Cluster Membership		
Case Number	Cluster	Distance
29	7	.281
30	10	.268
31	12	.032
32	12	.520
33	7	.155
34	1	.529
35	5	.426
36	12	.137
37	7	.086
38	10	.325
39	12	.295
40	9	.055
41	2	.029
42	3	.505

Table 6.12-Cluster membership for Solidity

Cluster Membership		
Case Number	Cluster	Distance
1	4	.007
2	8	.005
3	2	.003
4	9	.006
5	8	.009
6	6	.009
7	4	.010
8	8	.005
9	9	.002
10	8	.010
11	11	.003
12	8	.006
13	2	.009
14	8	.008

Cluster Membership		
Case Number	Cluster	Distance
15	2	.001
16	14	.007
17	14	.007
18	4	.003
19	9	.010
20	11	.009
21	11	.011
22	12	.000
23	10	.000
24	3	.003
25	8	.006
26	13	.000
27	8	.001
28	1	.002

Cluster Membership		
Case Number	Cluster	Distance
29	1	.002
30	2	.002
31	3	.004
32	3	.002
33	7	.000
34	3	.001
35	2	.005
36	11	.011
37	5	.000
38	6	.002
39	6	.007
40	11	.005
41	9	.001
42	9	.003

Table 6.13-Cluster membership for On Pixels

Cluster Membership		
Case Number	Cluster	Distance
1	4	2664.667
2	8	2026.250
3	2	1087.200
4	9	2399.800
5	8	3284.750
6	6	3376.333
7	4	3777.333
8	8	1670.250
9	9	902.800
10	8	3740.750
11	11	1240.000
12	8	2109.750
13	2	3286.800
14	8	2930.250

Cluster Membership		
Case Number	Cluster	Distance
15	2	418.800
16	14	2549.500
17	14	2549.500
18	4	1112.667
19	9	3831.200
20	11	3284.000
21	11	4251.000
22	12	.000
23	10	.000
24	3	927.333
25	8	2075.250
26	13	316.000
27	8	433.250
28	1	649.500

Cluster Membership		
Case Number	Cluster	Distance
29	1	649.500
30	2	931.200
31	3	1671.667
32	3	744.333
33	7	.000
34	13	316.000
35	2	1687.200
36	11	4187.000
37	5	.000
38	6	585.667
39	6	2790.667
40	11	1980.000
41	9	530.200
42	9	1058.800

Table 6.14-Cluster membership for Narrow-Factor

Cluster Membership		
Case Number	Cluster	Distance
1	2	.005
2	2	.005
3	3	.000
4	7	.003
5	5	.000
6	6	.000
7	7	.001
8	6	.001
9	9	.001
10	10	.000
11	11	.000
12	11	.001
13	8	.002
14	14	.005

Cluster Membership		
Case Number	Cluster	Distance
15	14	.008
16	8	.005
17	14	.003
18	8	.007
19	4	.001
20	1	.000
21	4	.001
22	12	.000
23	7	.003
24	6	.001
25	7	.001
26	7	.001
27	5	.000
28	7	.001

Cluster Membership		
Case Number	Cluster	Distance
29	7	.001
30	9	.001
31	7	.001
32	7	.004
33	7	.001
34	13	.006
35	13	.006
36	7	.001
37	7	.001
38	7	.001
39	7	.001
40	11	.001
41	11	.005
42	11	.003

After we obtain the clusters, it is to be judged that which feature bears the best cluster formation capability. Feature having the best cluster formation capability is the most important feature for the cluster analysis. K-Means clustering algorithm does not give this feature importance measure. Hence we discuss how this measure is obtained and compare in our experiment -

We have used total 15 different classes of tomato leaves and 14 classes of tomato fruits, with each class consisting of 3 cases (patterns). Now it is to be seen that how many cases of a particular class are included in a same cluster. More is the number of cases of a particular class included in same cluster, better is clustering result. We calculate a measure called "Same Cluster Membership Ratio (SCMR)" for each class from the cluster membership (Table 5 and Table 6) produced from a particular feature and then finding the summation of SCMRs (Total SCMR) of all classes. Higher is the value of Total SCMR of a feature, higher is its importance in cluster formation. SCMR for a particular class can be calculated by finding the ratio of the total number of cases of that class included in same cluster and total number of cases present in that class (in our experiment, this value is 3 for all classes) .

Now from the above definition of SCMR, it is obvious that in our problem, we will obtain one of the following three SCMR values for each class under the following conditions-

i) If no case belonging to a particular class is included in same cluster, then SCMR of that class is 0.

ii) If 2 cases of a class are included in same cluster, then SCMR of that class is 2/3=0.67(approx).

iii) If all 3 cases of a class are included in same cluster, then SCMR of that class is 3/3=1.

Thus, the range of values of SCMR in our problem is 0 to 1.

Table 7 and Table 8 show the class number and the corresponding three case numbers (pattern number) that belong to that particular class. Cluster number is the number of the cluster; the case pattern is a member of. And the last field of Table 7 and table 8 is the calculated SCMR value of each class for each feature variables of tomato leaf and fruit. The last row of each table shows the Total SCMR value related to each leaf/fruit feature---

Table 7-SCMR calculation for individual leaf features

SCMR calculation for Major Axis			
class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	4	
	5	5	
2	6	6	0.00
	7	7	
	8	15	
3	9	9	0.00
	10	10	
	11	7	
4	12	12	0.00
	13	15	
	14	7	
5	15	15	0.67
	16	10	
	17	15	
6	18	13	0.00
	19	15	
	20	12	
7	21	10	0.00
	22	8	
	23	10	
8	24	15	0.00
	25	15	
	26	11	
9	27	10	0.00
	28	8	
	29	7	
10	30	12	0.00
	31	11	
	32	13	
11	33	7	0.00
	34	12	
	35	7	
12	36	10	0.00
	37	10	
	38	15	
13	39	14	0.00
	40	3	
	41	12	
14	42	10	0.00
	43	8	
	44	15	
15	45	13	0.00
Total SCMR			0.67

SCMR calculation for Minor Axis			
class number	case number	cluster number	SCMR
	1	3	
	2	3	
1	3	3	1.00
	4	4	
	5	4	
2	6	3	0.67
	7	7	
	8	8	
3	9	7	0.67
	10	10	
	11	11	
4	12	12	0.00
	13	13	
	14	15	
5	15	15	0.67
	16	1	
	17	9	
6	18	14	0.00
	19	6	
	20	6	
7	21	9	0.67
	22	1	
	23	4	
8	24	3	0.00
	25	14	
	26	3	
9	27	2	0.00
	28	3	
	29	4	
10	30	1	0.00
	31	5	
	32	5	
11	33	5	1.00
	34	14	
	35	13	
12	36	15	0.00
	37	4	
	38	6	
13	39	9	0.00
	40	4	
	41	3	
14	42	4	0.67
	43	7	
	44	3	
15	45	6	0.00
Total SCMR			5.35

SCMR calculation for Aspect Ratio			
class number	case number	cluster number	SCMR
	1	15	
	2	8	
1	3	8	0.67
	4	11	
	5	8	
2	6	6	0.00
	7	6	
	8	8	
3	9	6	0.67
	10	8	
	11	11	
4	12	12	0.00
	13	6	
	14	1	
5	15	15	0.00
	16	3	
	17	7	
6	18	13	0.00
	19	3	
	20	5	
7	21	7	0.00
	22	3	
	23	2	
8	24	1	0.00
	25	14	
	26	14	
9	27	11	0.67
	28	14	
	29	9	
10	30	7	0.00
	31	6	
	32	1	
11	33	15	0.00
	34	1	
	35	15	
12	36	15	0.67
	37	13	
	38	5	
13	39	10	0.00
	40	6	
	41	13	
14	42	13	0.67
	43	8	
	44	1	
15	45	4	0.00
Total SCMR			3.35

SCMR calculation for Eccentricity			
class number	case number	cluster number	SCMR
	1	13	
	2	2	
1	3	8	0.00
	4	4	
	5	8	
2	6	6	0.00
	7	9	
	8	8	
3	9	9	0.67
	10	10	
	11	11	
4	12	12	0.00
	13	6	
	14	14	
5	15	13	0.00
	16	3	
	17	7	
6	18	5	0.00
	19	3	
	20	3	
7	21	7	0.67
	22	3	
	23	1	
8	24	14	0.00
	25	14	
	26	14	
9	27	4	0.67
	28	14	
	29	3	
10	30	7	0.00
	31	6	
	32	14	
11	33	13	0.00
	34	14	
	35	13	
12	36	13	0.67
	37	5	
	38	3	
13	39	15	0.00
	40	9	
	41	5	
14	42	1	0.00
	43	2	
	44	14	
15	45	7	0.00
Total SCMR			2.68

SCMR calculation for Area

class number	case number	cluster number	SCMR
	1	3	
	2	2	
1	3	3	0.67
	4	4	
	5	5	
2	6	6	0.00
	7	7	
	8	8	
3	9	9	0.00
	10	10	
	11	9	
4	12	12	0.00
	13	15	
	14	14	
5	15	15	0.67
	16	3	
	17	2	
6	18	15	0.00
	19	14	
	20	14	
7	21	3	0.67
	22	8	
	23	3	
8	24	11	0.00
	25	7	
	26	14	
9	27	1	0.00
	28	6	
	29	3	
10	30	3	0.67
	31	8	
	32	7	
11	33	15	0.00
	34	1	
	35	13	
12	36	11	0.00
	37	8	
	38	2	
13	39	2	0.67
	40	6	
	41	8	
14	42	11	0.00
	43	14	
	44	3	
15	45	8	0.00
Total SCMR			**3.35**

SCMR calculation for Diameter

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	2	
	5	5	
2	6	6	0.00
	7	11	
	8	8	
3	9	9	0.00
	10	7	
	11	11	
4	12	11	0.67
	13	13	
	14	14	
5	15	15	0.00
	16	7	
	17	7	
6	18	8	0.67
	19	14	
	20	11	
7	21	15	0.00
	22	7	
	23	6	
8	24	4	0.00
	25	15	
	26	6	
9	27	13	0.00
	28	6	
	29	13	
10	30	6	0.67
	31	6	
	32	12	
11	33	12	0.67
	34	13	
	35	13	
12	36	15	0.67
	37	14	
	38	13	
13	39	10	0.00
	40	4	
	41	7	
14	42	6	0.00
	43	6	
	44	13	
15	45	11	0.00
Total SCMR			**3.35**

SCMR calculation for Rectangularity

class number	case number	cluster number	SCMR
	1	5	
	2	2	
1	3	3	0.00
	4	4	
	5	4	
2	6	3	0.67
	7	7	
	8	1	
3	9	9	0.00
	10	10	
	11	11	
4	12	12	0.00
	13	13	
	14	14	
5	15	14	0.67
	16	3	
	17	3	
6	18	13	0.67
	19	5	
	20	5	
7	21	3	0.67
	22	3	
	23	5	
8	24	8	0.00
	25	8	
	26	6	
9	27	9	0.00
	28	5	
	29	6	
10	30	3	0.00
	31	8	
	32	1	
11	33	13	0.00
	34	14	
	35	1	
12	36	8	0.00
	37	6	
	38	5	
13	39	5	0.67
	40	2	
	41	15	
14	42	6	0.00
	43	14	
	44	15	
15	45	6	0.00
Total SCMR			**3.35**

SCMR calculation for Compactness

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	2	
	5	5	
2	6	12	0.00
	7	7	
	8	9	
3	9	9	0.67
	10	6	
	11	9	
4	12	9	0.67
	13	13	
	14	4	
5	15	15	0.00
	16	12	
	17	8	
6	18	9	0.00
	19	4	
	20	7	
7	21	4	0.67
	22	12	
	23	10	
8	24	14	0.00
	25	15	
	26	6	
9	27	13	0.00
	28	10	
	29	15	
10	30	6	0.00
	31	6	
	32	13	
11	33	7	0.00
	34	13	
	35	13	
12	36	15	0.67
	37	4	
	38	15	
13	39	9	0.00
	40	11	
	41	10	
14	42	6	0.00
	43	10	
	44	15	
15	45	7	0.00
Total SCMR			**2.68**

SCMR calculation for Perimeter Ratio of Major Axis-Minor Axis

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	5	
	5	5	
2	6	6	0.67
	7	7	
	8	8	
3	9	9	0.00
	10	9	
	11	11	
4	12	12	0.00
	13	1	
	14	14	
5	15	15	0.00
	16	7	
	17	14	
6	18	11	0.00
	19	7	
	20	9	
7	21	7	0.67
	22	1	
	23	14	
8	24	7	0.00
	25	14	
	26	13	
9	27	7	0.00
	28	4	
	29	13	
10	30	13	0.67
	31	14	
	32	13	
11	33	4	0.00
	34	8	
	35	14	
12	36	13	0.00
	37	15	
	38	8	
13	39	8	0.67
	40	6	
	41	15	
14	42	10	0.00
	43	14	
	44	13	
15	45	8	0.00
Total SCMR			**2.68**

SCMR calculation for Concavity

class number	case number	cluster number	SCMR
	1	3	
	2	2	
1	3	3	0.67
	4	4	
	5	5	
2	6	6	0.00
	7	7	
	8	8	
3	9	9	0.00
	10	10	
	11	9	
4	12	12	0.00
	13	15	
	14	14	
5	15	15	0.67
	16	3	
	17	2	
6	18	15	0.00
	19	14	
	20	14	
7	21	3	0.67
	22	8	
	23	3	
8	24	11	0.00
	25	7	
	26	14	
9	27	1	0.00
	28	6	
	29	3	
10	30	3	0.67
	31	8	
	32	7	
11	33	15	0.00
	34	1	
	35	13	
12	36	11	0.00
	37	8	
	38	2	
13	39	2	0.67
	40	6	
	41	8	
14	42	11	0.00
	43	14	
	44	3	
15	45	8	0.00
Total SCMR			**3.35**

SCMR calculation for Perimeter Ratio of Diameter

class number	case number	cluster number	SCMR
	1	3	
	2	2	
1	3	3	0.67
	4	4	
	5	5	
2	6	14	0.00
	7	1	
	8	8	
3	9	8	0.67
	10	10	
	11	11	
4	12	11	0.67
	13	14	
	14	14	
5	15	14	1.00
	16	1	
	17	6	
6	18	15	0.00
	19	9	
	20	15	
7	21	7	0.00
	22	10	
	23	1	
8	24	13	0.00
	25	6	
	26	10	
9	27	6	0.67
	28	9	
	29	6	
10	30	6	0.67
	31	10	
	32	14	
11	33	6	0.00
	34	9	
	35	10	
12	36	6	0.00
	37	10	
	38	7	
13	39	15	0.00
	40	13	
	41	12	
14	42	13	0.67
	43	14	
	44	1	
15	45	8	0.00
Total SCMR			**5.02**

SCMR calculation for R-Factor

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	4	
	5	5	
2	6	6	0.00
	7	11	
	8	9	
3	9	9	0.67
	10	10	
	11	11	
4	12	11	0.67
	13	13	
	14	14	
5	15	15	0.00
	16	10	
	17	7	
6	18	9	0.00
	19	14	
	20	11	
7	21	15	0.00
	22	7	
	23	6	
8	24	12	0.00
	25	15	
	26	6	
9	27	13	0.00
	28	6	
	29	13	
10	30	6	0.67
	31	10	
	32	11	
11	33	11	0.67
	34	13	
	35	13	
12	36	15	0.67
	37	14	
	38	13	
13	39	8	0.00
	40	12	
	41	10	
14	42	6	0.00
	43	10	
	44	13	
15	45	11	0.00
Total SCMR			**3.35**

Table 8-SCMR calculation for individual fruit features

SCMR calculation for Branch Length

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	4	
	5	4	
2	6	6	0.67
	7	9	
	8	8	
3	9	5	0.00
	10	8	
	11	11	
4	12	12	0.00
	13	5	
	14	9	
5	15	9	0.67
	16	11	
	17	8	
6	18	10	0.00
	19	11	
	20	5	
7	21	13	0.00
	22	9	
	23	14	
8	24	2	0.00
	25	5	
	26	8	
9	27	9	0.00
	28	14	
	29	6	
10	30	6	0.67
	31	14	
	32	4	
11	33	14	0.67
	34	13	
	35	9	
12	36	7	0.00
	37	2	
	38	4	
13	39	6	0.00
	40	9	
	41	9	
14	42	12	0.67
Total SCMR			**3.35**

SCMR calculation for Branch Width

class number	case number	cluster number	SCMR
	1	8	
	2	2	
1	3	3	0.00
	4	4	
	5	5	
2	6	14	0.00
	7	8	
	8	8	
3	9	9	0.67
	10	7	
	11	11	
4	12	9	0.00
	13	2	
	14	14	
5	15	2	0.67
	16	7	
	17	7	
6	18	1	0.67
	19	1	
	20	12	
7	21	5	0.00
	22	3	
	23	8	
8	24	9	0.00
	25	5	
	26	11	
9	27	12	0.00
	28	12	
	29	12	
10	30	9	0.67
	31	7	
	32	11	
11	33	11	0.67
	34	11	
	35	5	
12	36	7	0.00
	37	10	
	38	13	
13	39	6	0.00
	40	4	
	41	4	
14	42	14	0.67
Total SCMR			**4.02**

SCMR calculation for Length-Width Ratio

class number	case number	cluster number	SCMR
1	1	3	0.67
	2	2	
	3	3	
2	4	4	0.00
	5	2	
	6	6	
3	7	7	1.00
	8	7	
	9	7	
4	10	11	1.00
	11	11	
	12	11	
5	13	13	0.00
	14	14	
	15	12	
6	16	7	0.67
	17	11	
	18	11	
7	19	8	0.67
	20	12	
	21	8	
8	22	7	0.00
	23	11	
	24	8	
9	25	5	0.00
	26	8	
	27	12	
10	28	14	0.00
	29	4	
	30	5	
11	31	13	1.00
	32	13	
	33	13	
12	34	7	0.00
	35	5	
	36	3	
13	37	1	0.00
	38	9	
	39	10	
14	40	2	0.67
	41	2	
	42	4	
Total SCMR			**5.68**

SCMR calculation for Area

class number	case number	cluster number	SCMR
	1	1	
	2	8	
1	3	3	0.00
	4	9	
	5	5	
2	6	12	0.00
	7	7	
	8	12	
3	9	9	0.00
	10	10	
	11	9	
4	12	12	0.00
	13	9	
	14	3	
5	15	7	0.00
	16	12	
	17	6	
6	18	14	0.00
	19	2	
	20	13	
7	21	13	0.67
	22	7	
	23	11	
8	24	8	0.00
	25	14	
	26	7	
9	27	12	0.00
	28	7	
	29	7	
10	30	10	0.67
	31	12	
	32	12	
11	33	7	0.67
	34	11	
	35	4	
12	36	12	0.00
	37	7	
	38	10	
13	39	12	0.00
	40	8	
	41	8	
14	42	2	0.67
Total SCMR			**2.68**

SCMR calculation for Perimeter

class number	case number	cluster number	SCMR
	1	1	
	2	8	
1	3	3	0.00
	4	9	
	5	5	
2	6	12	0.00
	7	7	
	8	12	
3	9	9	0.00
	10	10	
	11	9	
4	12	12	0.00
	13	9	
	14	3	
5	15	7	0.00
	16	12	
	17	6	
6	18	14	0.00
	19	2	
	20	13	
7	21	13	0.67
	22	7	
	23	11	
8	24	8	0.00
	25	14	
	26	7	
9	27	12	0.00
	28	7	
	29	7	
10	30	10	0.67
	31	12	
	32	12	
11	33	7	0.67
	34	11	
	35	4	
12	36	12	0.00
	37	7	
	38	10	
13	39	12	0.00
	40	8	
	41	8	
14	42	2	0.67
Total SCMR			**2.68**

SCMR calculation for Equivalent Diameter

class number	case number	cluster number	SCMR
	1	4	
	2	8	
1	3	2	0.00
	4	9	
	5	8	
2	6	6	0.00
	7	4	
	8	8	
3	9	9	0.00
	10	8	
	11	11	
4	12	8	0.67
	13	2	
	14	8	
5	15	2	0.67
	16	14	
	17	14	
6	18	4	0.67
	19	9	
	20	11	
7	21	11	0.67
	22	12	
	23	10	
8	24	3	0.00
	25	8	
	26	13	
9	27	8	0.67
	28	1	
	29	1	
10	30	2	0.67
	31	3	
	32	3	
11	33	7	0.67
	34	13	
	35	2	
12	36	11	0.00
	37	5	
	38	6	
13	39	6	0.67
	40	11	
	41	9	
14	42	9	0.67
Total SCMR			**6.03**

SCMR calculation for Rectangularity

class number	case number	cluster number	SCMR
	1	2	
	2	2	
1	3	10	0.67
	4	6	
	5	9	
2	6	5	0.00
	7	12	
	8	7	
3	9	9	0.00
	10	10	
	11	6	
4	12	12	0.00
	13	2	
	14	5	
5	15	2	0.67
	16	10	
	17	7	
6	18	6	0.00
	19	6	
	20	5	
7	21	5	0.67
	22	14	
	23	13	
8	24	1	0.00
	25	2	
	26	11	
9	27	6	0.00
	28	3	
	29	3	
10	30	1	0.67
	31	11	
	32	8	
11	33	14	0.00
	34	14	
	35	9	
12	36	5	0.00
	37	4	
	38	4	
13	39	5	0.67
	40	5	
	41	5	
14	42	5	1.00
Total SCMR			**4.35**

SCMR calculation for Diameter

class number	case number	cluster number	SCMR
	1	1	
	2	3	
1	3	13	0.00
	4	10	
	5	5	
2	6	6	0.00
	7	9	
	8	7	
3	9	9	0.67
	10	5	
	11	9	
4	12	5	0.67
	13	5	
	14	10	
5	15	10	0.67
	16	5	
	17	10	
6	18	13	0.00
	19	3	
	20	11	
7	21	2	0.00
	22	4	
	23	10	
8	24	8	0.00
	25	4	
	26	7	
9	27	4	0.67
	28	6	
	29	12	
10	30	12	0.67
	31	6	
	32	10	
11	33	10	0.67
	34	9	
	35	5	
12	36	14	0.00
	37	5	
	38	10	
13	39	6	0.00
	40	13	
	41	13	
14	42	5	0.67
Total SCMR			**4.69**

SCMR calculation for Perimeter Ratio of Branch Length-Branch Width

class number	case number	clusternumber	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	10	
	5	13	
2	6	13	0.67
	7	7	
	8	7	
3	9	7	1.00
	10	10	
	11	10	
4	12	6	0.67
	13	13	
	14	14	
5	15	13	0.67
	16	13	
	17	11	
6	18	5	0.00
	19	7	
	20	6	
7	21	6	0.67
	22	7	
	23	12	
8	24	4	0.00
	25	9	
	26	3	
9	27	9	0.67
	28	3	
	29	6	
10	30	2	0.00
	31	2	
	32	7	
11	33	2	0.67
	34	8	
	35	5	
12	36	5	0.67
	37	9	
	38	11	
13	39	13	0.00
	40	6	
	41	3	
14	42	3	0.67
Total SCMR			**6.36**

SCMR calculation for Convexity

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	14	0.00
	4	4	
	5	5	
2	6	12	0.00
	7	7	
	8	12	
3	9	4	0.00
	10	10	
	11	4	
4	12	12	0.00
	13	4	
	14	14	
5	15	7	0.00
	16	12	
	17	6	
6	18	11	0.00
	19	3	
	20	8	
7	21	8	0.67
	22	7	
	23	13	
8	24	9	0.00
	25	11	
	26	7	
9	27	12	0.00
	28	7	
	29	7	
10	30	10	0.67
	31	12	
	32	12	
11	33	7	0.67
	34	1	
	35	5	
12	36	12	0.00
	37	7	
	38	10	
13	39	12	0.00
	40	9	
	41	2	
14	42	3	0.00
Total SCMR			**2.01**

SCMR calculation for Perimeter Ratio of Diameter

class number	case number	cluster number	SCMR
	1	1	
	2	2	
1	3	3	0.00
	4	6	
	5	2	
2	6	6	0.67
	7	11	
	8	2	
3	9	9	0.00
	10	9	
	11	11	
4	12	4	0.00
	13	6	
	14	13	
5	15	6	0.67
	16	4	
	17	5	
6	18	2	0.00
	19	2	
	20	2	
7	21	7	0.67
	22	2	
	23	8	
8	24	10	0.00
	25	5	
	26	9	
9	27	7	0.00
	28	13	
	29	13	
10	30	6	0.67
	31	6	
	32	6	
11	33	6	1.00
	34	12	
	35	2	
12	36	14	0.00
	37	3	
	38	3	
13	39	13	0.67
	40	13	
	41	13	
14	42	10	0.67
Total SCMR			**5.02**

SCMR calculation for Solidity

class number	case number	cluster number	SCMR
	1	4	
	2	8	
1	3	2	0.00
	4	9	
	5	8	
2	6	6	0.00
	7	4	
	8	8	
3	9	9	0.00
	10	8	
	11	11	
4	12	8	0.67
	13	2	
	14	8	
5	15	2	0.67
	16	14	
	17	14	
6	18	4	0.67
	19	9	
	20	11	
7	21	11	0.67
	22	12	
	23	10	
8	24	3	0.00
	25	3	
	26	13	
9	27	8	0.00
	28	1	
	29	1	
10	30	2	0.67
	31	3	
	32	3	
11	33	7	0.67
	34	3	
	35	2	
12	36	11	0.00
	37	5	
	38	6	
13	39	6	0.67
	40	11	
	41	9	
14	42	9	0.67
Total SCMR			**5.36**

SCMR calculation for On Pixels			
class number	case number	cluster number	SCMR
	1	4	
	2	8	
1	3	2	0.00
	4	9	
	5	8	
2	6	6	0.00
	7	4	
	8	8	
3	9	9	0.00
	10	8	
	11	11	
4	12	8	0.67
	13	2	
	14	8	
5	15	2	0.67
	16	14	
	17	14	
6	18	4	0.67
	19	9	
	20	11	
7	21	11	0.67
	22	12	
	23	10	
8	24	3	0.00
	25	8	
	26	13	
9	27	8	0.67
	28	1	
	29	1	
10	30	2	0.67
	31	3	
	32	3	
11	33	7	0.67
	34	13	
	35	2	
12	36	11	0.00
	37	5	
	38	6	
13	39	6	0.67
	40	11	
	41	9	
14	42	9	0.67
Total SCMR			6.03

SCMR calculation for Narrow-Factor			
class number	case number	cluster number	SCMR
	1	2	
	2	2	
1	3	3	0.67
	4	7	
	5	5	
2	6	6	0.00
	7	7	
	8	6	
3	9	9	0.00
	10	10	
	11	11	
4	12	11	0.67
	13	8	
	14	14	
5	15	14	0.67
	16	8	
	17	14	
6	18	8	0.67
	19	4	
	20	1	
7	21	4	0.67
	22	12	
	23	7	
8	24	6	0.00
	25	7	
	26	7	
9	27	5	0.67
	28	7	
	29	7	
10	30	9	0.67
	31	7	
	32	7	
11	33	7	1.00
	34	13	
	35	13	
12	36	7	0.67
	37	7	
	38	7	
13	39	7	1.00
	40	11	
	41	11	
14	42	11	1.00
Total SCMR			8.36

Total SCMR value of each of the leaf features (Table 9) and fruit features (Table 10) defines the fact that 'Minor Axis'(Total SCMR value 5.35) ' and 'Major Axis'(Total SCMR value 0.67) are the leaf features with highest and lowest Total SCMR value. Hence 'Minor Axis' and 'Major - Axis' are the most important and least important leaf features in terms of cluster formation respectively. Where as fruit feature having the highest and lowest importance are 'Narrow-Factor'(Total SCMR value 8.36, highest among all fruit features) and 'Convexity' (Total SCMR value 2.01, lowest among all fruit features) respectively.

Table 9- Leaf features with their Total SCMR values

FEATURE NAME	TOTAL SCMR
Major Axis	0.67
Minor Axis	5.35
Aspect Ratio	3.35
Eccentricity	2.68
Area	3.35
Rectangularity	3.35
Diameter	3.35
Compactness	2.68
Perimeter Ratio of Major Axis-Minor Axis	2.68
Perimeter Ratio of Diameter	5.02
Concavity	3.35
R-Factor	3.35

Table 10- Fruit features with their Total SCMR values

FEATURE NAME	TOTAL SCMR
Branch Length	3.35
Branch Width	4.02
Length Width Ratio	5.68
Area	2.68
Perimeter	2.68
Equivalent Diameter	6.03
Rectangularity	4.35
Diameter	4.69
Perimeter Ratio of Branch Length-Branch Width	6.36
Perimeter Ratio of Diameter	5.02
Convexity	2.01
Solidity	5.36
On Pixels	6.03
Narrow-Factor	8.36

4.2 Two-step Clustering Results

Two-step clustering algorithm has used 12 and 14 number of feature variables of tomato leaf and fruit as input and 15 & 14 final clusters of leaf and fruit are produced (Figure 1).

Model Summary

Algorithm	TwoStep
Inputs	12
Clusters	15

Model Summary

Algorithm	TwoStep
Inputs	14
Clusters	14

Figure 1: Summary of Two-step clustering

Figure 2 shows the quality of the cluster formation of tomato leaf and fruit based on Silhouette measure of cohesion and separation. Silhouette measure is obtained with in the range of -1.0 to 1.0. Silhouette measure of greater than 0.5 signifies that the cluster formation is of a good quality. So from Figure 2, it is obvious that Two-step clustering has formed good quality clusters from sample patterns of tomato leaf and fruit.

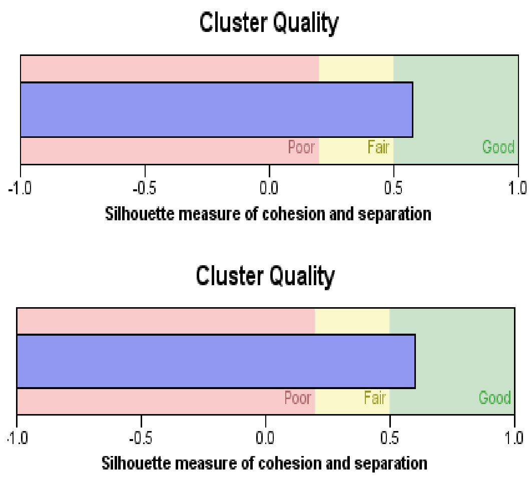

Figure 2: Cluster formation quality for leaf and fruit.

Number of patterns placed in a cluster defines the size of that cluster. Figure 3 depicts the pie chart representation of the leaf and fruit clusters according to their sizes expressed in terms of percentage of the total number of leaf and patterns.

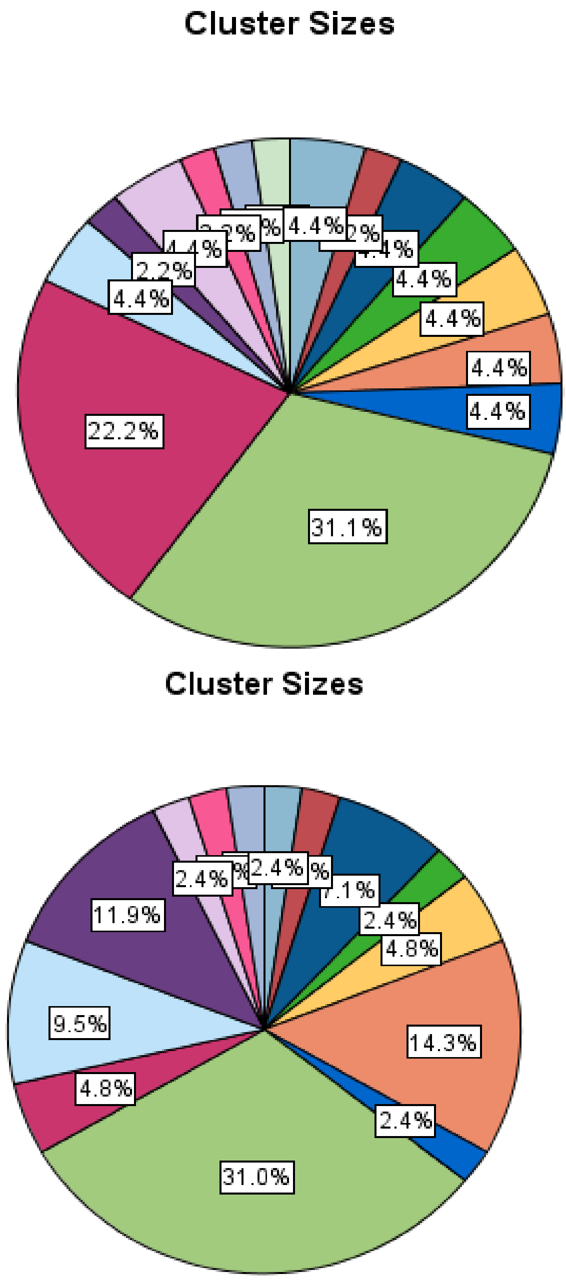

Figure 3: Pie chart of leaf and fruit clusters (size wise)

All features don't contribute in the same manner in the cluster formation process. Some features contribute highly where contribution of some features may be low. Features with higher level of contribution have higher importance in the formation of clusters. Figure 4 shows the top-down arrangement of leaf and fruit features in descending order of their importance in cluster formation. According to Figure 4, Compactness and Major Axis are the most important and least important leaf features. Where as Rectangularity and Branch Length are the fruit features with highest and lowest importance.

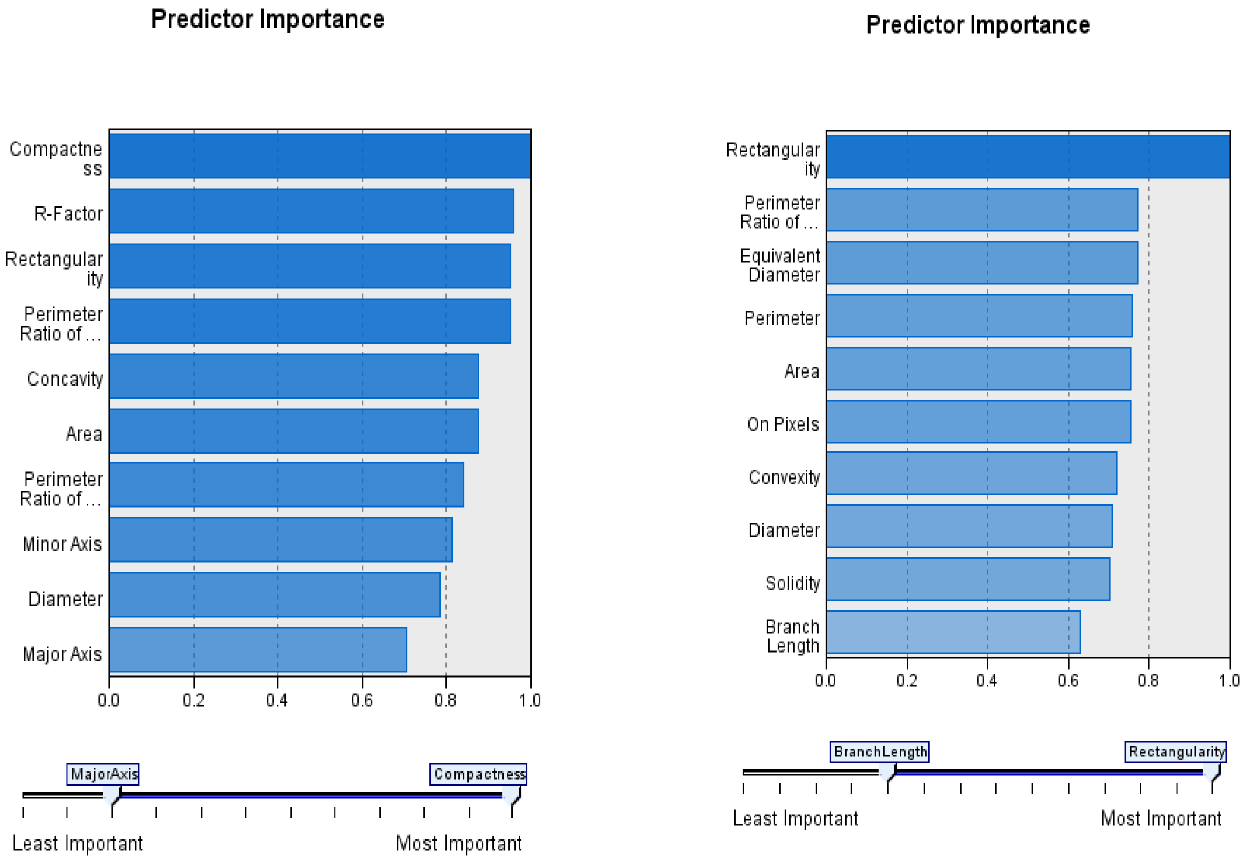

Figure 4: Order of Importance of leaf and fruit features

5. CONCLUSION

The scheme for K-Means and Two-step clustering algorithm to discriminate the tomato leaf, fruiting habit image samples along with morphological feature importance has been introduced. Formation of valid clusters assures the successful execution of the clustering techniques on the feature set and there by reflecting the categorical distribution of tomato species. Feature importance calculation through SCMR measure declares Minor Axis and Narrow-Factor as the respective leaf and fruit features upon which best cluster formation is observed. Where as according to Two-step clustering, Compactness and Rectangularity are the most important leaf and fruit feature as per as cluster formation capability is concerned. Considering this phenomena, one of the aspects of future work is to use these features to validate a large volume image dataset of tomato leaves and fruits. Another important futuristic aspect is to build up a leaf / fruit categorization system with relevant feedback mechanisms to help the persons related to cultivation process and collection of feedback from them for the enhancement of the system. Also applying the other renowned clustering methods like fuzzy clustering, neural network based clustering on the sample data set and hence analyzing the result is a future work.

6. REFERENCES

[1] A.Hazra, K.Deb, S.Kundu, P.Hazra,"Shape Oriented Feature Selection for Tomato Plant Identification", International Journal of Computer Applications Technology and Research,Volume 2-Issue 4, 2013, pages 449-454.

[2] S.Kundu, A.Hazra, K.Deb, P.Hazra, "Dimensionality Reduction of Morphological features of Tomato Leaves and Fruiting Habits", IEEE International Conference on Communications, Devices and Intelligent Systems(CODIS 2012),pages 608-611

[3] Gregory A. Wilkin, Xiuzhen Huang,"K-Means Clustering Algorithms: Implementation and Comparison", IEEE Second International Multisymposium on Computer and Computational Sciences, 2007, pages 133-136.

[4] Li-Qing Li, Qiao Liu, Han-qing Zhou,"Research on Patient Satisfaction Degree Evaluation of Three A-level Hospital in Jiangxi Province Based on Cluster Analysis", IEEE International Conference on Information Management, Innovation Management and Industrial Engineering, 2011, pages 563-567.

[5] Jie Yao, "Research on the Application of K-means cluster analysis in undergraduate instructional management", IEEE International Conference on Advanced Computer Control, 2008, pages 628-631.

[6] Bin Lie, Huichao Zhang, Huiyu Chen, Lili Liu, Dingwei Wang, "A K-means Clustering Based Algorithm for Shill Bidding Recognition in Online Auction", IEEE 24th Chinese Control and Decision Conference(CCDC 2012), pages 939-943.

[7] Shuhua Ren, Alin Fan, "*K*-means Clustering Algorithm Based On Coefficient of Variation", IEEE 4th International Congress on Image and Signal Processing, 2011, pages 2076-2079.

[8] M.Narasimha Murty, V.Susheela Devi, "Pattern Recognition An Algorithmic Approach Undergraduate Topics in Computer Science",Springer 2011, Universities Press (India) Pvt.Ltd. 207-229

[9] Andrew R. Webb, "Statistical Pattern Recognition", Second Edition, 2002, John Wiley & Sons,Ltd, 361-402