The Pharma Innovation Journal 2022; 11(2): 1112-1118

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; 11(2): 1112-1118

www.thepharmajournal.com
Received: 02-11-2021
Accepted: 13-01-2022

SK Mundotiya
M.Sc. Scholar, Division of Seed Science & Technology, ICAR-IARI, New Delhi, India

Jagdish Goyanka
Ph.D scholar, Division of NBPGR, ICAR- IARI, New Delhi, India

Ravi Kumar Mundotiya
M.sc in Biotechnology
Dr. B.L. Institute of Technology Jaipur, Rajasthan, India

Sunil Ramling Swami
ARS, ICAR-IGFRI Jhansi, Uttar Pradesh, India

Corresponding Author:
SK Mundotiya
M.Sc. Scholar, Division of Seed Science & Technology, ICAR-IARI, New Delhi, India

Cultivar differentiation in rice's extant varieties through image analysis

SK Mundotiya, Jagdish Goyanka, Ravi Kumar Mundotiya and Sunil Ramling Swami

Abstract
There are about 950 released and notified varieties of rice (*Oryza sativa* L.) in India for which diagnostic features are well known and the same are followed for the purpose of seed certification. Hence, variety identification is of prime importance. The scope of morphological differences between the varieties is less due to narrow genetic base, and requires skilled human power which is subjective in nature. Also this process is time, labour and cost intensive. Keeping in mind the above facts, the present study was initiated with the objective of making differentiation in rice varieties using image analysis technique. The experimental material comprised of twenty eight extant rice varieties. A complete digital database comprising of 84 images each for seed and 448 leaf images were generated. Two different types of softwares were used for extraction of features from the images *viz*. Grain Analysis Software (for size and shape features) and MATLAB software (for textural features). The varieties were grouped on the basis of these features generated from seed and leaf images. Seed imaging features differentiated the varieties into six clusters whereas leaf imaging features alone and combination of seed and leaf images grouped the varieties into seven clusters. The study revealed that image features extracted from seed were most helpful for distinguishing the varieties because it is supported by two kind of study, one based on Grain Analysis Software and second based on MATLAB software whereas study of leaf for cultivar differentiation is based on MATLAB software only.

Keywords: Image analysis, MATLAB software, cultivar differentiation, clustering, etc.

Introduction
Rice is the world's most important food crop and provides 21% of the human per capita energy, and 15% of protein globally (FAO). There are about 950 released and notified varieties of rice (*Oryza sativa* L.) in India for which diagnostic features are well known and the same are followed for the purpose of seed certification. Hence, variety identification is of prime importance (Chakrabarty *et al.* 2012) [1]. The scope of morphological differences between the varieties is less due to narrow genetic base, and requires skilled human power which is subjective in nature. Also this process is time, labour and cost intensive. Hence, the manual identification of varieties/seeds by specialized technicians is slow, has low reproducibility, and possesses a degree of subjectivity that is hard to quantify. Implication of new techniques for addressing a particular variety can be focused and attention is being laid at international level for the development of suitable lab techniques like image analysis of seed or plant organs, bio chemical and molecular markers. Seed size, shape and colour of nine Italian cultivars of common vetch (*Vicia sativa* L.) and one of hairy vetch (*Vicia villosa* Roth.) were measured using digital images acquired by a flatbed scanner and, on the basis of a Linear Discriminant Analysis algorithm, a statistical classifier was able to identify the ten cultivars (Grillo *et.al* 2011) [2]. Geetha *et al.* (2011) [3] studied geometry of single seed in mustard with image analysis technique. A two-stage classifier combining distance discriminate and a back propagation neural network (BPNN) was built for identification. On the first stage, corn kernels were divided into three types: white, yellow, and mixed corn by distance discriminate analysis. And then different varieties in the same type were identified by an improved BPNN classifier (Chen xiao *et al.*, 2009) [4]. Zayas *et al.* (1989) [5] used image analysis to discriminate between wheat and non-wheat and between weed seeds and stones in the non-wheat part of grain sample. Keeping in mind the above facts, the present study was initiated with the objective of making differentiation in rice varieties using image analysis technique.

Materials and Methods
The experimental material comprised of twenty eight extant rice varieties which are sown during kharif season of 2014-15 in nursery field of Seed Science and Technology Division, and Genetic Division, IARI New Delhi. The laboratory experiments were carried out in the Division of Seed Science and Technology, Indian Agriculture Research Institute New Delhi.

Imaging Protocol

The imaging of plant parts (leaf and seed) was done by the following ways:

Seed	**Scanner** 1. Canon LiDE 110 2. Resolution: 600 Dpi 3. Document size: Platen
Leaf	**Photography set up** 1. Height of Camera: 18cm 2. Lens focus setting: between 0.3m-0.5m 3. Lighting type: Back light with standard reflector 4. Light power setting: 1.0 5. Distance of light source: 20cm below as per size 6. Exposure parameter: F=20,ISO=100,Shutter speed =60, White balance= Auto flesh 7. Gadget used for holding sample: Cartridge sheet

Image Processing and extraction of features:

The processing of images was done at Seed Science and Technology Division, IARI New Delhi and CIAE, Bhopal MP. The image processing was done by using two different kinds of softwares:

1. Grain Analysis Software developed by Dr. Nachiket Kotwaliwale.
   - Used for extraction of size and shape features from the seed.
2. MATLAB software (version 7.12.0.635,R2011a) developed by Dr. Nachiket.
   - Used for extraction of textural features from the seed and leaf images.

Eleven parameters are measured from the Grain Analysis Software software: Area, Perimeter, Bounding box length, Bounding box width, Axial length, Axial width, Median length, Median width, Eccentricity, Roundness, Equivalent diameter.

The basic features recorded by the MATLAB software:

1. **Morphological features:** Length, Width, Awn length, Kernel area, Kernel perimeter, Major axis, Minor axis, Eccentricity, Equivalent Diameter, Length-width ratio.
2. **Textural features:** Contrast; Correlation; Energy; Homogeneity; Range; STD; Entropy; Offset 0; Offset 45; Offset 90; Offset 135; SRE; LRE; GLN; LP; RLN; LGRE; HGRE.
3. **Chromatic features:** Redness; Greenness; Blueness; Hue; Saturation; Value; Hue Std; RHS colour value.

Result and Discussion

For seed imaging, flat bed scanner Canon LiDE 110 version 1.2.00 was used to develop the image library. Ten rice grains for each variety were placed on scanner avoiding grain to grain contact. The images were taken in three replications per variety. That accounted for a total of 28 x 3 = 84 images; with an average of 10 seeds per image. The seed data generated by the Grain Analysis Software is presented in Table 1. Table 1 shows the average results for samples of 10 grains of one cultivar, computed over three replications, for all the 28 varieties. The data for three replications showed comparable values for all the parameters. This demonstrated the accuracy of the scanner for digital inputs of the image attributes. From the values of axis length, perimeter and area in the data generated by the Grain Analysis Software, five shape factors were derived. The different shape factors showed that varieties T. Basmati and Vivek Dhan 62 were significantly different from each other for Shape factor 1, varieties Vasumati and PR 113 for Shape factor 2, varieties PB 1509 and Vivek Dhan 62 for Shape factors 3 and 4, and T. Basmati and Jaya were significantly different from each other for Shape factor 5. A comparative assessment of various shape factors revealed that shape factor 5 involving area and major and minor axis length was most useful for distinguishing varieties.

Textural features extracted from Seed images and Leaf images

In addition to Grain Analysis Software, MATLAB software was used to extract morphological, textural and chromatic features which are mentioned in material and method part, from seed and leaf separately. For leaf imaging, two type of leaves were imaged i.e. flag leaf and penultimate leaf (leaf next to flag leaf) using photography set-up as given in Materials and Methods. In both leaves, both the sides of leaves i.e. ventral (upper) and dorsal (lower) sides were imaged. Hence, the sample size consisted of four images for each kind of leaf per side i.e flag leaf ventral, flag leaf dorsal, penultimate leaf ventral and penultimate leaf dorsal. Thus, total number of images generated were (4+4+4+4) x 28 varieties = 448 images. The 28 varieties were found to differ significantly from each other with respect to these textural features. Since these features were exclusively related to the seed and leaf images, hence an attempt was made to classify/ group the varieties on the basis of features extracted from the images.

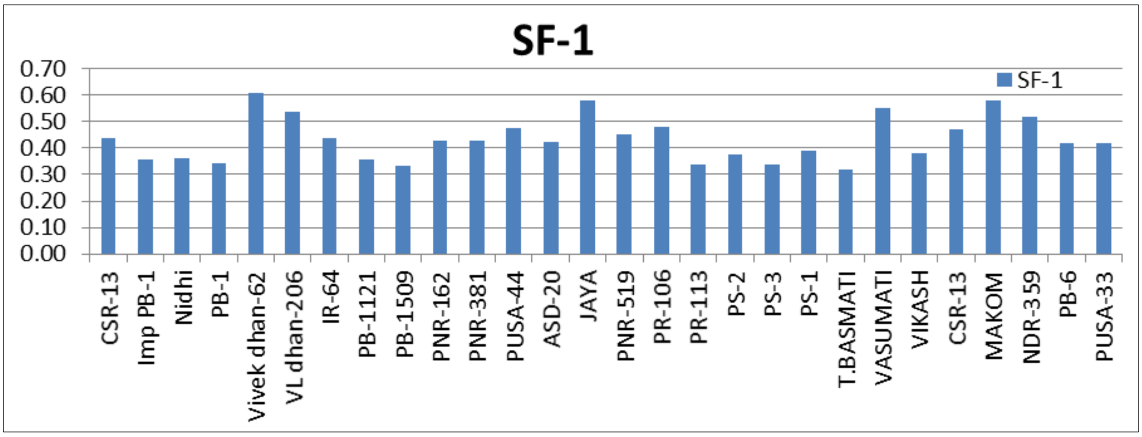
Table 1: Seed data generated by Grain analysis software

S. No	VAR	Area, mm^2	Length, mm	Breadth, mm	Eccentricity	Perimeter, mm	Equivalent Dia, mm	Roundness	Axial Length, mm	Axial Width, mm	Median Length, mm	Median Width, mm
1	CSR-13	13.421	8.74	2.1	0.973	19.66	4.13	0.2227	4.8	1.98	3.5221	1.66
2	Imp PB-1	19.912	11.93	2.49	0.98	26.4	5.04	0.1782	6.35	2.26	5.0165	1.9
3	Nidhi	14.278	10.1	2.06	0.979	22.3	4.26	0.1791	5.03	1.83	4.0131	1.56
4	PB-1	17.636	11.55	2.17	0.982	25.44	4.73	0.1678	7.18	1.97	4.8301	1.66
5	Vivek dhan-62	15.768	7.71	2.79	0.931	18.07	4.48	0.3386	3.94	2.64	2.7646	2.31
6	VL dhan-206	17.09	8.74	2.69	0.948	20.01	4.66	0.2867	3.74	2.58	2.6946	2.19
7	IR-64	15.538	9.41	2.33	0.97	21.16	4.42	0.223	7.37	2.09	3.9475	1.81
8	PB-1121	21.639	12.65	2.39	0.985	27.6	5.25	0.1719	5.61	2.31	3.2006	1.85
9	PB-1509	21.783	13.25	2.31	0.984	28.76	5.24	0.1592	7.82	2.23	4.6059	1.88
10	PNR-162	13.658	8.92	2.16	0.971	19.97	4.18	0.2199	4.32	2	3.086	1.66
11	PNR-381	14.776	9.36	2.13	0.975	20.86	4.34	0.2156	4.13	2.04	3.1412	1.7
12	PUSA-44	16.009	9.05	2.34	0.967	20.58	4.5	0.249	5.59	2.22	4.1274	1.93
13	ASD-20	16.574	9.82	2.37	0.973	22.12	4.6	0.2197	5.63	2.12	4.6057	1.85

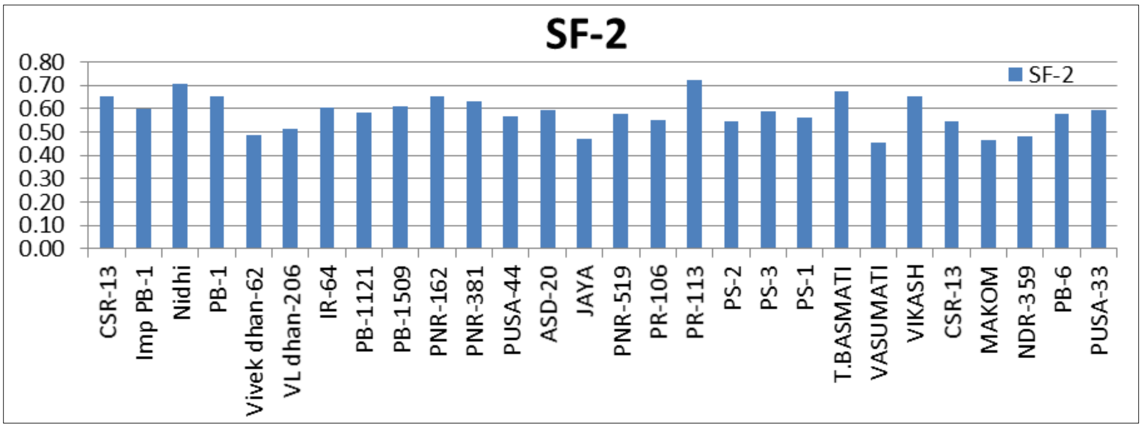
14	JAYA	17.848	8.42	2.91	0.937	19.68	4.75	0.3214	4.12	2.8	2.9294	2.38
15	PNR-519	16.108	9.33	2.32	0.968	21.13	4.52	0.236	7.36	2.17	4.6354	1.94
16	PR-106	16.64	9.15	2.53	0.961	20.87	4.6	0.2544	6.28	2.36	3.884	2
17	PR-113	14.678	10.6	2.09	0.981	23.43	4.3	0.1663	6.49	1.86	5.4567	1.51
18	PS-2	22.663	12.38	2.61	0.978	27.53	5.37	0.1896	10.05	2.37	5.3637	2.08
19	PS-3	21.907	12.9	2.52	0.984	28.52	5.26	0.1679	9.66	2.23	5.9139	1.92
20	PS-1	20.841	11.75	2.41	0.977	25.96	5.15	0.1928	6.39	2.23	5.2747	2.01
21	T.BASMATI	17.584	11.88	2.42	0.981	26.33	4.72	0.1587	8.77	2.13	4.119	1.67
22	VASUMATI	20.312	9.29	2.9	0.949	21.53	5.08	0.2989	5.31	2.77	3.7762	2.42
23	VIKASH	15.836	10.31	2.31	0.975	22.89	4.48	0.1893	5.6	2.05	4.2122	1.71
24	CSR-13	17.601	9.58	2.47	0.966	21.69	4.73	0.2436	5.63	2.32	4.5847	2.02
25	MAKOM	18.23	8.52	2.84	0.94	19.91	4.81	0.3202	4.66	2.72	3.3974	2.39
26	NDR-359	19.732	9.54	2.78	0.955	21.88	5.01	0.2769	4.68	2.65	3.6577	2.28
27	PB-6	17.464	10.13	2.34	0.974	22.9	4.71	0.216	6.66	2.2	5.1096	1.89
28	PUSA-33	17.173	10.2	2.35	0.974	22.67	4.68	0.2112	5.29	2.18	3.446	1.85

Shape Factors based on Feature extraction of Grains

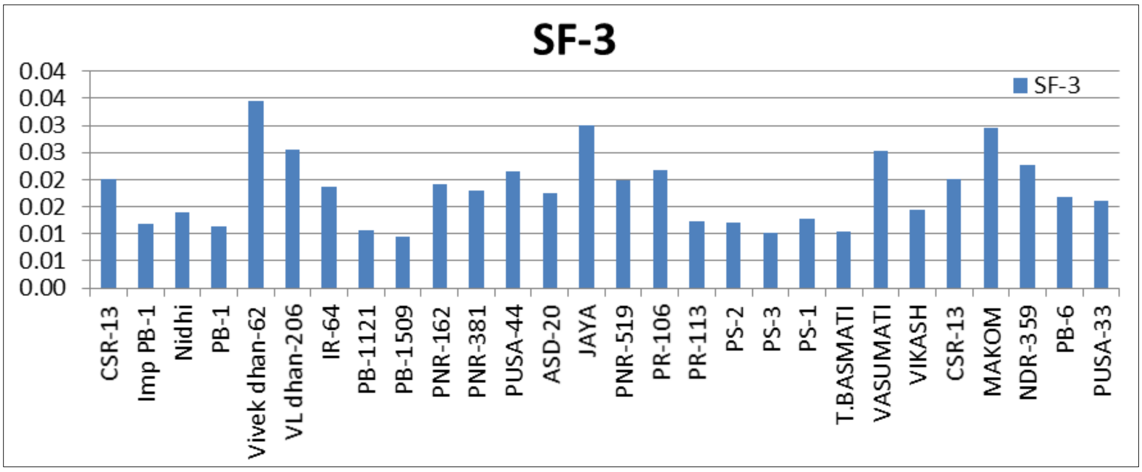
From the values of axis length, perimeter and area, shape factors were derived, following Symons and Fulcher (1988) formulae.



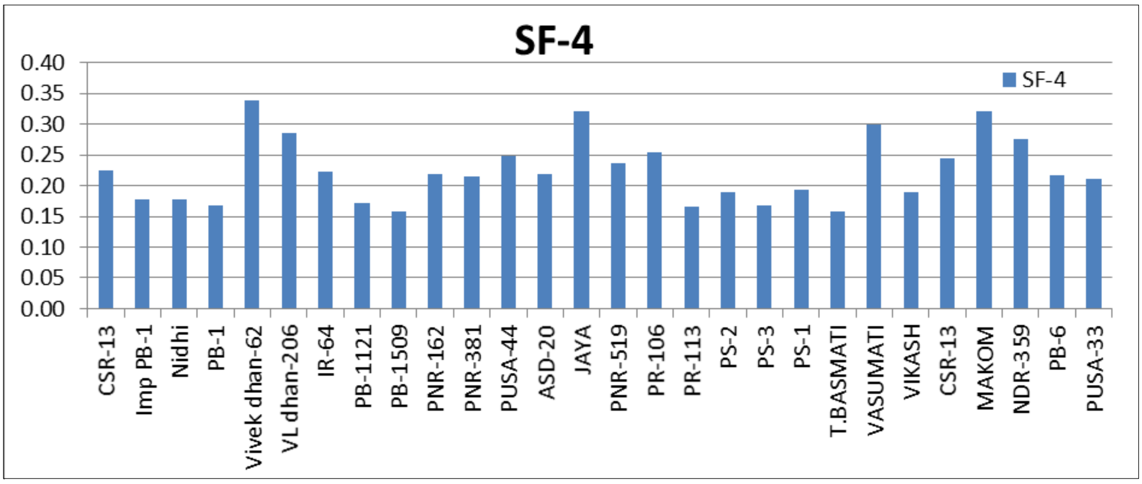
Shape factor 1: 4 π Area/Perimeter^2



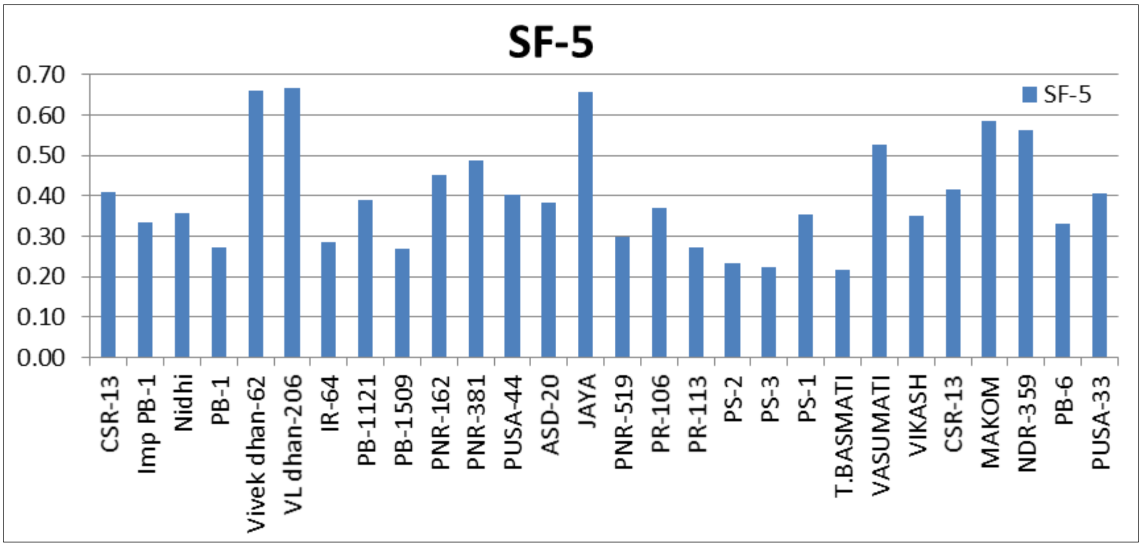
Shape factor 2: Major axis length/Area



Shape factor 3: Area/Major axis length^3



Shape factor 4: Area/(Major axis length/2)(Major axis length/2)



Shape factor 5: Area/(Major axis length/2)(Minor axis length/2)

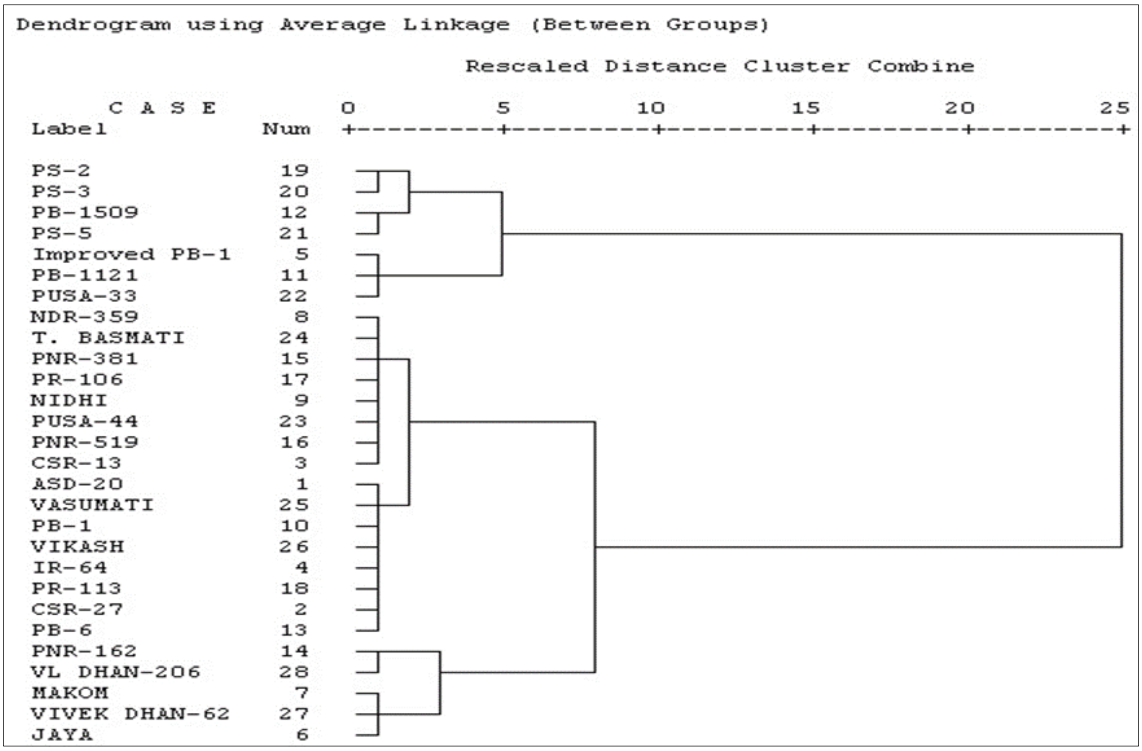


Fig 1: Dendrogram generation based on Seed data generated through MATLAB softwar

Table 2: Clustering pattern of varieties based on Seed data generated through MATLAB software

Cluster	Number of varieties	Variety Name
I	4	PS-2, PS-3, PB-1509, PS-5,
II	3	IMPROVED PB-1, PB-1121, PUSA-33
III	8	NDR-359, TARAORI BASMATI, PNR-381, PR-106, NIDHI, PUSA-44, PNR-519, CSR-13
IV	8	ASD-20, VASUMATI, PB-1, VIKASH, IR-64, PR-113, CSR-27, PB-6
V	2	PNR-162, VL DHAN-206,
VI	3	MAKOM, VIVEK DHAN-62, JAYA

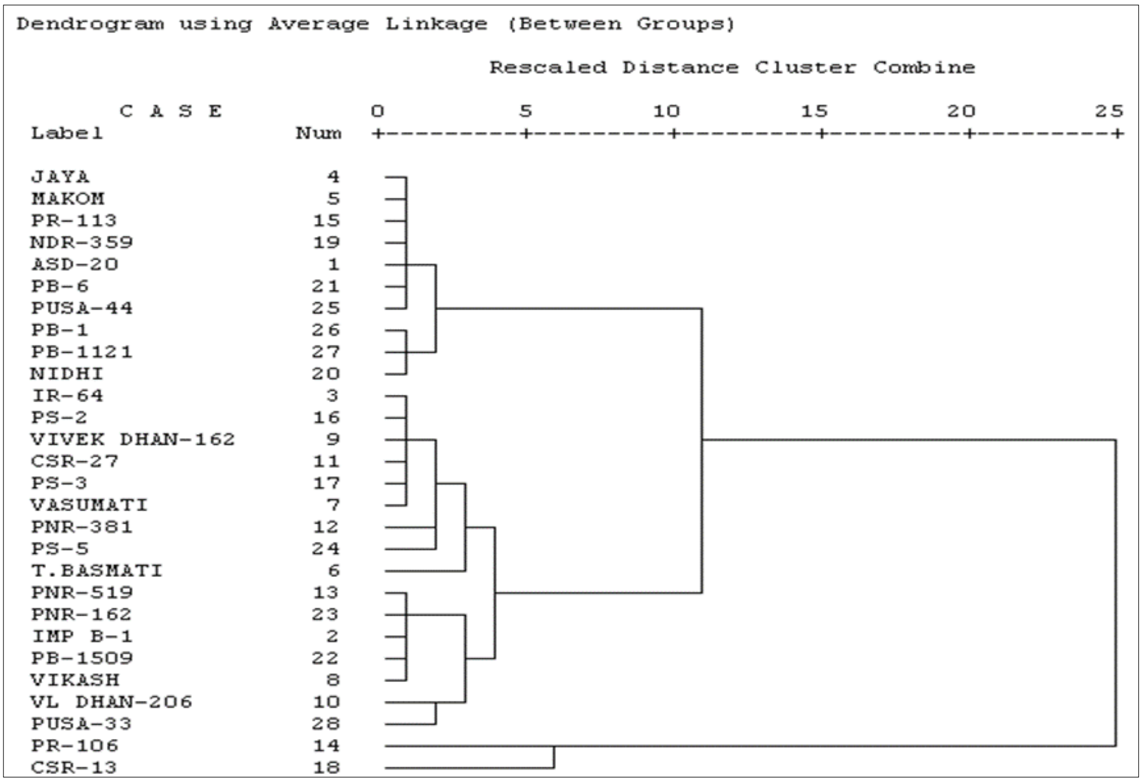


Fig 2: Dendrogram generation based on Leaves data generated through MATLAB softwar

Table 3: Clustering pattern of varieties based on Leaf data

Cluster	Number of varieties	Variety Name
I	7	JAYA, MAKOM PR-113, NDR-359, ASD-20, PB-6, PUSA-44
II	3	PB-1, PB-1121, NIDHI
III	6	IR-64, PS-2, VIVEK DHAN-62, CSR-27, PS-3, VASUMATI
IV	3	PNR-381, PS-5, TARAORI BASMATI
V	5	PNR-519, PNR-162, IMPROVED PB-1, PB-1509, VIKASH,
VI	2	VLDHAN-206, PUSA-33
VII	2	CSR-13, PR-106

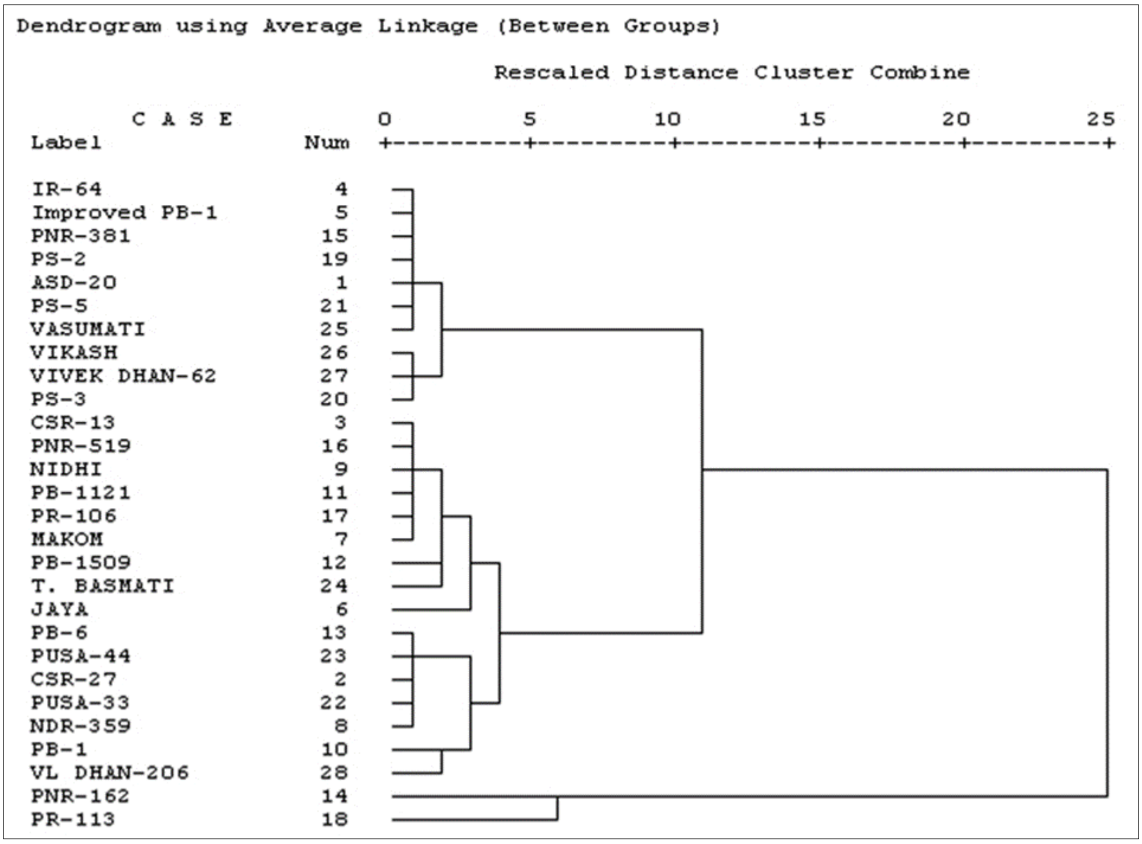


Fig 3: Dendrogram generation based on Seed and Leaves data

Table 4: Clustering pattern of varieties based on Seed and Leaves data

Cluster	Number of varieties	Variety Name
I	7	IR-64, IMPROVED PB-1, PNR-381, PS-2, ASD-20, PS-5, VASUMATI,
II	3	VIKASH, VIVEK DHAN-62, PS-3
III	6	CSR-13, PNR-519, NIDHI, PB-1121, PR-106, MAKOM
IV	3	PB-1509, TARAORI BASMATI, JAYA
V	5	PB-6, PUSA-44, CSR-27, PUSA-33, NDR-359,
VI	2	PB-1, VLDHAN-206
VII	2	PR-113, PNR-162

The dendrograms were generated from textural data related to seed (Fig.1), leaf (Fig.2) and combination of both seed and leaf (Fig.3). On the basis of dendrogram, cluster analysis was also performed for seed (Table 2), leaf (Table 3) and combination of seed and leaf (Table 4).

Since the objective of study is cultivar differentiation in Rice's extant varieties through image analysis, the cluster analysis based on textural features extracted by MATLAB software divided the 28 rice varieties into six different clusters on seed basis, into seven different clusters each on the basis of leaf and combination of seed and leaf as well. Based on the comparative clustering patterns on three different kinds of features *viz*. seed imaging, leaf imaging and seed imaging in combination with leaf imaging; the study revealed that image features extracted from seed were most helpful for distinguishing the varieties because it is supported by two kind of study based on Grain Analysis Software and MATLAB software whereas study of leaf for cultivar differentiation is based on MATLAB software only.

Ropelewska, E., & Rutkowski, K. P. (2021) [6]. Evaluated the usefulness of individual parts of fruit (skin, flesh, stone and seed) for cultivar discrimination of peaches based on textures determined using image analysis. Discriminant analysis was performed using the classifiers of Bayes net, logistic, SMO, multi-class classifier and random forest based on a set of combined textures selected from all color channels R, G, B, L, a, b, X, Y, Z and for textures selected separately for RGB, Lab and XYZ color spaces. In the case of sets of textures selected from all color channels (R, G, B, L, a, b, X, Y, Z), the accuracy of 100% was observed for flesh, stones and seeds for selected classifiers. Seed morphological features obtained by processing radiographs with the Tomato Analyzer software and of red–green–blue obtained and processed on the Groundeye device were used to test differentiation of materials of the *Urochloa* genus. Seeds of *Urochloa brizantha*, *Urochloa ruziziensis* and *Urochloa decumbens* were evaluated. Morphological features obtained by Tomato Analyzer allowed differentiation of *Urochloa* seeds at an accuracy level greater than 80% for all the materials evaluated De Freitas *et al.* (2021) [7].

Principal component analysis (PCA), linear discrimination analysis (LDA), partial least squares discriminant analysis (PLSDA), AdaBoost and support vector machine (SVM) methods were applied to classify seeds of sweet clover and alfalfa according to their morphological features and spectral traits or a combination thereof. The results showed that an excellent classification could be achieved based on a combination of morphological features and spectral data in a tested data set Hu, X *et al*.,(2020) [8]. To develop discriminative models based on geometric features to distinguish seeds belonging to different apple cultivars, images of seeds of apples 'Gala', 'Jonagold' and 'Idared' were acquired using a flatbed scanner. In the case of models build based on selected linear dimensions, the accuracy of discrimination was equal up to 84% for distinguishing seeds of all three apple cultivars for the J48 classifier from Decision Trees and 93% for analysis of 'Gala' and 'Idared' for the J48 from Decision Trees. (Ropelewska, E., & Rutkowski, K. P. 2021) [9].

Pacifico, L *et al.* (2019) [10] develop a new medicinal plant data set based on the extraction of texture and color features from plant leaf images. A complete automatic plant recognition system is proposed, and five well-known machine learning classifiers are tested as the recognition module. Experimental results showed that the best classifiers are able to obtain average accuracies over 97% on the proposed data set. Vasanthan, V *et al.*, (2019) [11] used digital image analysis for identification and discrimination of crop varieties in Sesamum crop and found that Cluster analysis revealed that the varieties could be grouped into two major clusters in which CO 1, TMV 3, TMV 4, TMV 5, TMV 7 formed one cluster whereas the other varieties were grouped under another cluster, which showed that the genotypes in one cluster had similarity in most of the parameters and also its parentage. Thus, image analysis helps in discriminating the morphological variation in seeds related to genotype and its evolution. Pereira, C. S *et al.* (2018) [12] proposed a segmentation algorithm based on region growing using color model and threshold techniques for classification of the pixels belonging to vine leaves from vineyard color images captured in real field environment. Concerning boundary-based measures of quality, an average accuracy of 94.8% over a 140 image dataset was achieved. It proves that the proposed method gives suitable results for an ongoing research work for automatic identification and characterization of different endogenous grape varieties of the Portuguese Douro Demarcated Region.

Reference

1. Chakrabarty SK, Joshi MA, Singh Y, Maity A, Vashisht V, Dadlani M. Characterization and evaluation of variability in farmers' varieties of rice from West Bengal. Indian Journal of Genetics and Plant Breeding (The). 2012;72(2):136-142.
2. Grillo O, Miceli C, Venora G. Computerised image analysis applied to inspection of vetch seeds for varietal identification. Seed Science and Technology. 2011;39(2):490-500.
3. Geetha VV, Balamurugan P, Bhaskaran M. Characterization of mustard genotypes through image analysis. Research Journal of Seed Science. 2011;4(4):192-198.
4. Chen Xiao, Li Xunyi, Wei Zhang Junxiong. Combining discriminate analysis and neural network for corn variety identification. Computers and electronic in agriculture. 2009;71S:S48-S53.
5. Zayas I, Pomeranz Y, Lai FS. Discrimination of wheat and nonwheat components in grain samples by image analysis. Cereal Chemistry. 1989;66(3):233-237.
6. Ropelewska E, Rutkowski KP. Differentiation of peach cultivars by image analysis based on the skin, flesh, stone and seed textures. European Food Research and Technology. 2021, 1-7.
7. De Freitas MN, Dias MAN, Gomes-Junior FG, Abud HF, de Araújo LB, de Moraes TF. Discrimination of Urochloa seed genotypes through image analysis: Morphological features. Agronomy Journal. 2021;113(6):4930-4944.
8. Hu X, Yang L, Zhang Z, Wang Y. Differentiation of alfalfa and sweet clover seeds via multispectral imaging. Seed Science and Technology. 2020;48(1):83-99.
9. Ropelewska E, Rutkowski KP. Cultivar discrimination of stored apple seeds based on geometric features determined using image analysis. Journal of Stored Products Research. 2021;92:101804.
10. Pacifico LD, Britto LF, Oliveira EG, Ludermir TB. Automatic classification of medicinal plant species based on color and texture features. In 2019 8th Brazilian Conference on Intelligent Systems (BRACIS). IEEE. 2019 Oct, 741-746.
11. Vasanthan V, Geetha R, Menaka C, Vakeswaran V, Chidambaram K. Characterization of sesame varieties through image analysis. Electronic Journal of Plant Breeding. 2019;10(2):785-790.
12. Pereira CS, Morais R, Reis MJ. Pixel-based leaf segmentation from natural vineyard images using color model and threshold techniques. In International Conference Image Analysis and Recognition. Springer, Cham. 2018 June, 96-106.