

The data set (and description) can be downloaded here:

<http://archive.ics.uci.edu/ml/machine-learning-databases/image/>

**Description:**

1. Title: Image Segmentation data

2. Source Information

- Creators: Vision Group, University of Massachusetts
- Donor: Vision Group (Carla Brodley, brodley@cs.umass.edu)
- Date: November, 1990

3. Past Usage: None yet published

4. Relevant Information:

The instances were drawn randomly from a database of 7 outdoor images. The images were handsegmented to create a classification for every pixel.

Each instance is a 3x3 region.

5. Number of Instances: Training data: 210 Test data: 2100

6. Number of Attributes: 19 continuous attributes

## 7. Attribute Information:

1. region-centroid-col: the column of the center pixel of the region.
2. region-centroid-row: the row of the center pixel of the region.
3. region-pixel-count: the number of pixels in a region = 9.
4. short-line-density-5: the results of a line extractoin algorithm that counts how many lines of length 5 (any orientation) with low contrast, less than or equal to 5, go through the region.
5. short-line-density-2: same as short-line-density-5 but counts lines of high contrast, greater than 5.
6. vedge-mean: measure the contrast of horizontally adjacent pixels in the region. There are 6, the mean and standard deviation are given. This attribute is used as a vertical edge detector.
7. vegde-sd: (see 6)
8. hedge-mean: measures the contrast of vertically adjacent pixels. Used for horizontal line detection.
9. hedge-sd: (see 8).
10. intensity-mean: the average over the region of  $(R + G + B)/3$
11. rawred-mean: the average over the region of the R value.
12. rawblue-mean: the average over the region of the B value.
13. rawgreen-mean: the average over the region of the G value.
14. exred-mean: measure the excess red:  $(2R - (G + B))$
15. exblue-mean: measure the excess blue:  $(2B - (G + R))$
16. exgreen-mean: measure the excess green:  $(2G - (R + B))$
17. value-mean: 3-d nonlinear transformation of RGB. (Algorithm can be found in Foley and VanDam, Fundamentals of Interactive Computer Graphics)
18. saturatoin-mean: (see 17)
19. hue-mean: (see 17)

## 8. Missing Attribute Values: None

## 9. Class Distribution:

Classes: brickface, sky, foliage, cement, window, path, grass.

30 instances per class for training data.

300 instances per class for test data.

## Citation Request:

Please refer to the repository <http://archive.ics.uci.edu/ml> (see citation policy).

See also Frank, A. & Asuncion, A. (2010). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.

**Descriptive statistics:**

Dataset= segmentation : n= 660 , d= 10

Class1: n= 330

**Covariance matrix:**

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]
[1,]	4618.9952	223.3175	-0.3150	-0.0296	14.3268	246.3601	45.4397	-240.4204	-1.0481	1.9278
[2,]	223.3175	1428.3118	0.0574	-0.0817	12.3275	40.9520	6.2346	-305.3255	1.1930	-0.1304
[3,]	-0.3150	0.0574	0.0021	0.0001	-0.0192	-0.0741	-0.0132	0.0228	0.0000	0.0002
[4,]	-0.0296	-0.0817	0.0001	0.0005	0.0010	-0.0047	0.0064	0.0017	0.0000	0.0000
[5,]	14.3268	12.3275	-0.0192	0.0010	12.6602	48.3600	-0.5104	-6.2857	-0.0040	-0.0211
[6,]	246.3601	40.9520	-0.0741	-0.0047	48.3600	911.8912	10.7377	-3.0201	-0.0896	0.0114
[7,]	45.4397	6.2346	-0.0132	0.0064	-0.5104	10.7377	12.3857	-3.3884	-0.0462	0.0558
[8,]	-240.4204	-305.3255	0.0228	0.0017	-6.2857	-3.0201	-3.3884	240.9849	-0.6699	-0.1435
[9,]	-1.0481	1.1930	0.0000	0.0000	-0.0040	-0.0896	-0.0462	-0.6699	0.0062	-0.0055
[10,]	1.9278	-0.1304	0.0002	0.0000	-0.0211	0.0114	0.0558	-0.1435	-0.0055	0.0136

**Correlation matrix:**

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]
[1,]	1.0000	0.0869	-0.1019	-0.0192	0.0592	0.1200	0.1900	-0.2279	-0.1952	0.2436
[2,]	0.0869	1.0000	0.0334	-0.0954	0.0917	0.0359	0.0469	-0.5204	0.3996	-0.0296
[3,]	-0.1019	0.0334	1.0000	0.1224	-0.1187	-0.0539	-0.0827	0.0323	0.0033	0.0407
[4,]	-0.0192	-0.0954	0.1224	1.0000	0.0121	-0.0068	0.0805	0.0047	-0.0094	-0.0138
[5,]	0.0592	0.0917	-0.1187	0.0121	1.0000	0.4501	-0.0408	-0.1138	-0.0142	-0.0509
[6,]	0.1200	0.0359	-0.0539	-0.0068	0.4501	1.0000	0.1010	-0.0064	-0.0375	0.0032
[7,]	0.1900	0.0469	-0.0827	0.0805	-0.0408	0.1010	1.0000	-0.0620	-0.1661	0.1361
[8,]	-0.2279	-0.5204	0.0323	0.0047	-0.1138	-0.0064	-0.0620	1.0000	-0.5462	-0.0794
[9,]	-0.1952	0.3996	0.0033	-0.0094	-0.0142	-0.0375	-0.1661	-0.5462	1.0000	-0.5932
[10,]	0.2436	-0.0296	0.0407	-0.0138	-0.0509	0.0032	0.1361	-0.0794	-0.5932	1.0000

Median: 135.2506 100.2022 0.0184 0.0022 3.0362 3.2383 2.6272 45.8064 0.2962 -2.0162

Mean: 130.9576 98.3909 0.0182 0.0034 2.9882 5.7268 2.548 44.8697 0.3087 -2.0328

**MCD-estimated:**

MDC-0.975-Mean: 122.1807 91.4463 0 0 2.1212 1.4752 1.3522 48.7548 0.3095 -2.0631

MDC-0.750-Mean: 122.3408 91.375 0 0 2.0956 1.4755 1.3542 48.7151 0.3095 -2.063

MDC-0.500-Mean: 122.1807 91.4463 0 0 2.1212 1.4752 1.3522 48.7548 0.3095 -2.0631

Class2: n= 330

Covariance matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]
[1,]	4404.4038	-578.3487	-0.0136	0.1671	3.3112	-25.2207	-11.7488	-29.2875	-3.4030	6.9387
[2,]	-578.3487	1172.6146	0.0088	0.0105	-3.5759	27.2982	-2.8230	-106.1587	-0.4449	6.2493
[3,]	-0.0136	0.0088	0.0006	0.0000	0.0030	-0.0016	0.0031	0.0388	-0.0001	-0.0010
[4,]	0.1671	0.0105	0.0000	0.0004	0.0140	0.0592	0.0090	0.0122	0.0004	-0.0012
[5,]	3.3112	-3.5759	0.0030	0.0140	3.3612	9.8051	0.7588	4.3410	0.0311	-0.3178
[6,]	-25.2207	27.2982	-0.0016	0.0592	9.8051	94.9681	7.0006	6.4833	0.1727	-0.4547
[7,]	-11.7488	-2.8230	0.0031	0.0090	0.7588	7.0006	3.5637	3.9886	0.0347	-0.2933
[8,]	-29.2875	-106.1587	0.0388	0.0122	4.3410	6.4833	3.9886	81.8782	-0.2237	-2.8163
[9,]	-3.4030	-0.4449	-0.0001	0.0004	0.0311	0.1727	0.0347	-0.2237	0.0806	-0.1285
[10,]	6.9387	6.2493	-0.0010	-0.0012	-0.3178	-0.4547	-0.2933	-2.8163	-0.1285	0.4921

Correlation matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]
[1,]	1.0000	-0.2545	-0.0083	0.1260	0.0272	-0.0390	-0.0938	-0.0488	-0.1806	0.1490
[2,]	-0.2545	1.0000	0.0104	0.0154	-0.0570	0.0818	-0.0437	-0.3426	-0.0458	0.2602
[3,]	-0.0083	0.0104	1.0000	0.0380	0.0653	-0.0065	0.0669	0.1735	-0.0102	-0.0566
[4,]	0.1260	0.0154	0.0380	1.0000	0.3821	0.3042	0.2381	0.0672	0.0645	-0.0886
[5,]	0.0272	-0.0570	0.0653	0.3821	1.0000	0.5488	0.2192	0.2617	0.0597	-0.2471
[6,]	-0.0390	0.0818	-0.0065	0.3042	0.5488	1.0000	0.3805	0.0735	0.0624	-0.0665
[7,]	-0.0938	-0.0437	0.0669	0.2381	0.2192	0.3805	1.0000	0.2335	0.0647	-0.2215
[8,]	-0.0488	-0.3426	0.1735	0.0672	0.2617	0.0735	0.2335	1.0000	-0.0871	-0.4437
[9,]	-0.1806	-0.0458	-0.0102	0.0645	0.0597	0.0624	0.0647	-0.0871	1.0000	-0.6451
[10,]	0.1490	0.2602	-0.0566	-0.0886	-0.2471	-0.0665	-0.2215	-0.4437	-0.6451	1.0000

Median: 171.9031 115.4398 0.0045 0.0037 1.1552 1.4977 1.0339 9.5113 0.4952 -1.7887

Mean: 160.0212 112.5303 0.0051 0.0037 1.1926 2.0731 1.0835 8.8438 0.5102 -1.8096

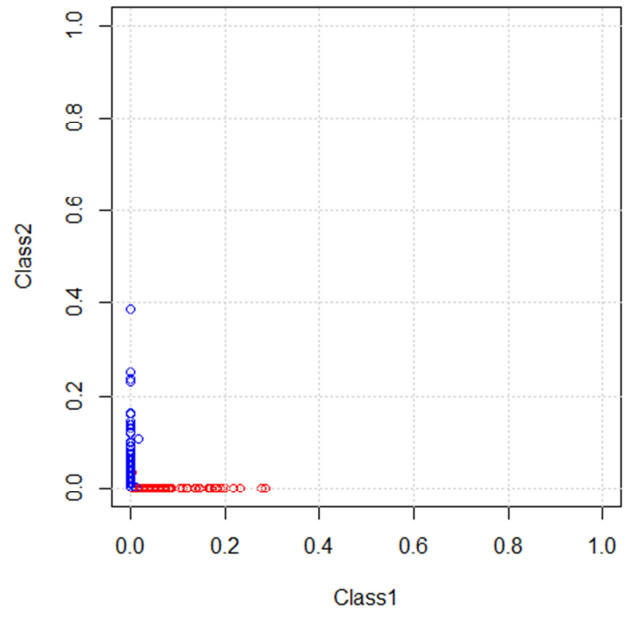
MCD-estimated:

MDC-0.975-Mean: 155.4624 112.919 0 0 0.5527 0.357 0.4785 8.1871 0.6401 -2.068  
MDC-0.750-Mean: 152.3837 109.2034 0 0 0.6731 0.4482 0.5268 8.8607 0.6734 -2.0674  
MDC-0.500-Mean: 154.2441 113.6394 0 0 0.5685 0.3695 0.4509 8.3159 0.639 -2.0695

Measures:

Mah.Dist: 3.1768  
Mah.Dist-MCD-0.975: 3.771  
Mah.Dist-MCD-0.750: 3.6156  
Mah.Dist-MCD-0.500: 3.6156

**DD-Plot (zonoid): segmentation**



**DD-Plot (random Tukey): segmentation**

