

Research Report

Submitted to :
Professor David A Landgrebe
&
Mr. Larry Biehl

Submitted by:
Eddy Dharma

Date:
07/06/1998

Content Page:

1) Introduction

- Aim
- Procedure
- Results
- Discussion

2) Appendix A

- Classified image taken from an image flown over West Lafayette County.
- An example of image generated from the Matlab subroutine (each row corresponds to each class).

3) Appendix B

- Results generated from Multispec Application Program.

Introduction:

Through the advancement of technology, the sensor now available contains as many as 220 channels. These vast amounts of channels are ideal in theoretical point of view due to the fact that many types of classes are now separable in this high dimensional space. In our case there are 220 dimensions in which all the distribution of the classes reside in. In theory, since high dimensional space contains mostly of empty space and hence no two pixels will lie in the same space, it will then be perfectly separable if and only if we could produce ample training samples to draw the boundaries between the classes of interest.

In practice, we might not be able to produce as many training samples so as to distinguish the different distributions of the classes in this high dimensional space. This poses a problem.

There are many ways to overcome this problem. One of them is the Projection Pursuit algorithm. This algorithm projects the high dimensional space into a smaller dimensional space and yet retaining the separability between classes. Furthermore, the distributions of classes in this smaller dimension resembles more towards a normally distributed function.

Other ways consists of DAFE (Discriminant Analysis Feature extraction) and DBFE (decision boundary feature extraction).

These 2 algorithms also involve in the reduction of features (a feature will consist of a linear combination of the 220 channels) from 220 features to, for example, 10 features, for a given image with 9 classes. The choice of how many features will have to be determined from the number of classes and also from the eigenvalues obtained from the transformation matrix when one runs the DAFE or DBFE algorithm.

Of course, once one has determined the ways to reduce the original high dimensions into a lower dimensions, the use of ECHO, Extraction and Classification of Homogeneous Object, (uses not only the spectral information but also the spatial information in the image) classifier could further improve the classification accuracy. Enhance Statistics is also widely used as a mean of utilizing the unlabeled samples coupled with the training information so as to hopefully gives a better classification accuracy.

Aim:

However, in this research, the objective is to look at the Leave-One-Out statistics using the Quadratic classifier. In theory, as few as 3 samples are capable of classifying a given class of interest. However, is it a valid assumption? The question of interest also lies in whether we could come out

with a rough estimation of a minimum number of samples in relation with the number of original number of features or dimensions so as to be able to give a fair amount of classification accuracy. That is if we have $n=220$ channels, we want to know $x \cdot n$ amount of training samples that will give us a reasonable classification accuracy.

Procedure:

Using an Aviris image of the West Lafayette area, first we choose a subset of the image and perform selection of the training samples for each class.

Having done that, the next step is to select at random a subset of the training samples one has chosen and run the classification for at least 10 times. How this is done will be briefly discussed as follow. For each of the training fields that we have selected, load that portion of the image (which is the training field) and do that for all fields pertaining to that particular class. Repeat the procedure for all classes in the subset of image that we have chosen to do the experiment on.

Up to this point, each class's training fields will be in a form of matrix with the following characteristic:

$X(1,:)$ is the pixel from row 1, column 1 in the image

$X(2,:)$ is the pixel from row 1, column 2 in the image and so on.

Next, we write a matlab subroutine to arbitrarily choose a subset of the matrices for each class. For example, 90%, 80% and so on. Having done that, we save all these matrices into a multispectral image. Now each row of this new image formed belong to the types of classes which we have defined.

Proceed with classification of the image with the new image generated as the training samples. Note that each row of the new image generated from the Matlab subroutine comprise of each class's training pixels. So in essence, we are classifying the image with this new image that we have created.

Results:

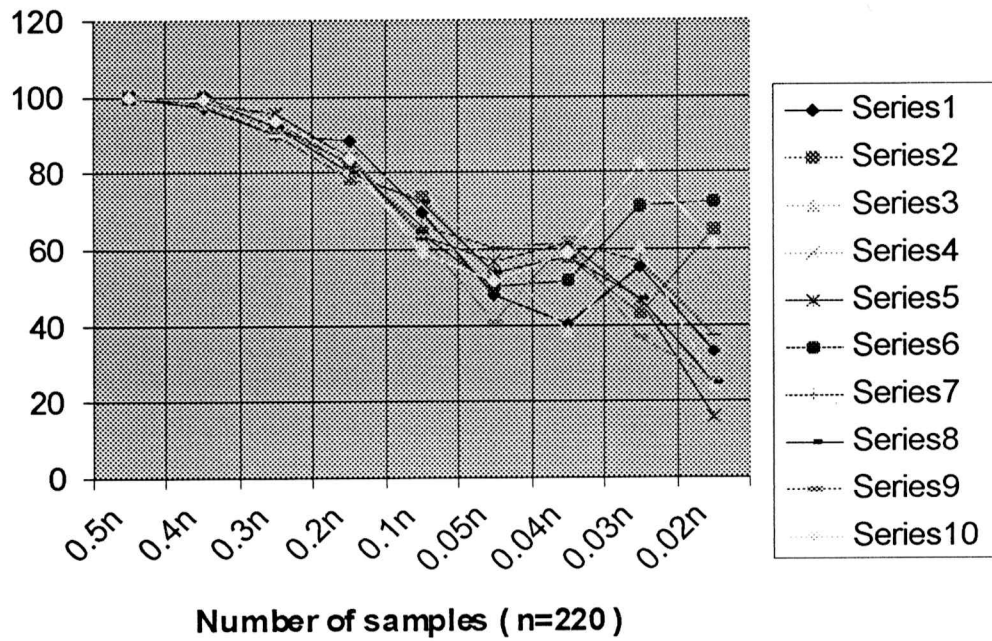
1) Building

This class has 112 samples altogether. This is approximately $0.5n$.

The following table shows the experimental results.

Number of samples		0.5n	0.4n	0.3n	0.2n	0.1n	0.05n	0.04n	0.03n	0.02n
trial	1	100	97.3	90.2	88.4	69.6	48.2	40.2	55.4	33
trial	2	100	98.2	90.2	78.6	73.2	53.6	58	42.9	65.2
trial	3	100	98.2	91.1	81.2	61.6	58	38.4	60.7	24.1
trial	4	100	99.1	94.6	79.5	78.6	54.5	58.9	45.5	53.6
trial	5	100	98.2	95.5	83.9	63.4	57.1	61.6	46.4	16.1
trial	6	100	100	92	83	64.3	50	51.8	71.4	72.3
Trial	7	100	100	93.8	82.1	66.1	60.7	61.6	57.1	37.5
Trial	8	100	100	92	80.4	72.3	53.6	58	47.3	25
Trial	9	100	99.1	91.1	84.8	63.4	41.1	62.5	36.6	25.9
Trial	10	100	99.1	93.8	83.9	58.9	51.8	58.9	82.1	61.6

**Plot of training accuracy for Building
against the number of samples**



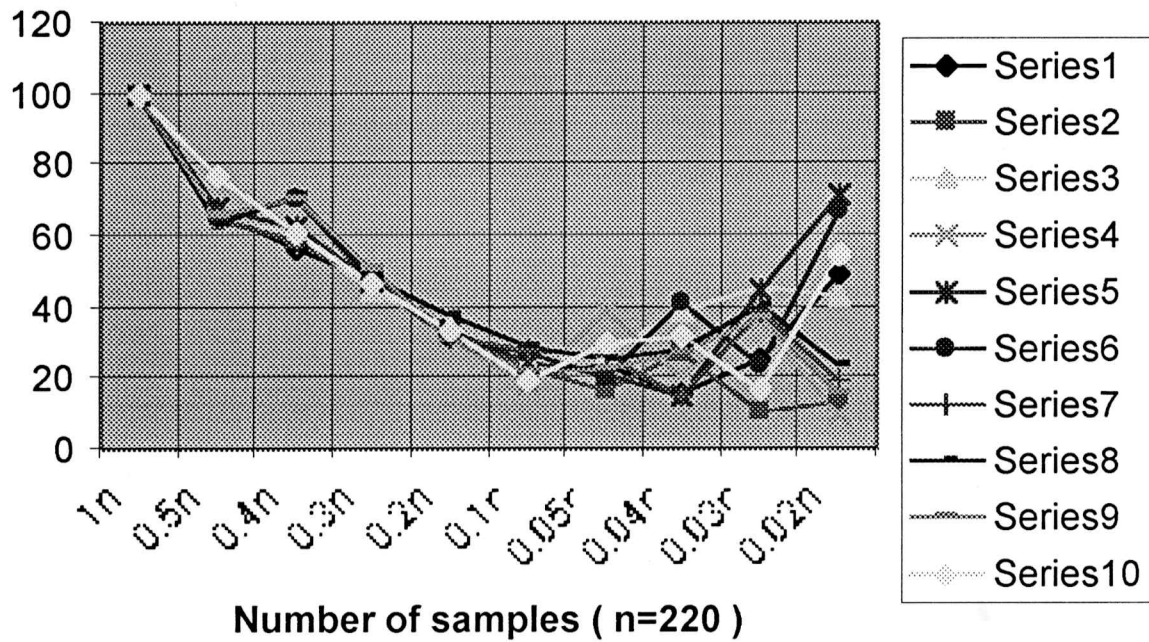
2) CornmintilleW

This class has 220 samples altogether. This is equal to $1n$.

The following table shows the experimental results:

Number of samples		$1n$	$0.5n$	$0.4n$	$0.3n$	$0.2n$	$0.1n$	$0.05n$	$0.04n$	$0.03n$	$0.02n$
trial	1	100	67.3	56.8	48.2	32.3	21.4	24.1	15.5	25.5	49.5
trial	2	99.5	64.5	59.1	46.8	33.6	22.7	16.8	27.3	10.5	13.6
trial	3	99.5	65.9	60.9	45	34.1	24.5	27.7	38.6	45.9	43.2
trial	4	99.5	65	62.3	48.2	31.8	18.2	39.1	20.5	35.9	9.5
trial	5	99.5	68.2	62.3	47.3	33.2	25	20.9	15	45	71.4
trial	6	99.5	65.5	70.9	47.7	31.8	27.7	20	41.8	23.6	66.8
trial	7	99.5	65	58.6	48.6	31.4	28.2	26.8	15	39.1	19.5
trial	8	99.5	63.6	71.4	48.6	37.7	29.5	25.5	28.2	40.9	23.6
trial	9	99.5	65.9	70.9	50.9	30.5	21.8	23.2	27.7	37.3	15.9
trial	10	99.5	76.4	60.5	46.8	33.2	19.1	30	31.4	16.8	54.5

Plot of training accuracy for CornmintilIEW against the number of samples



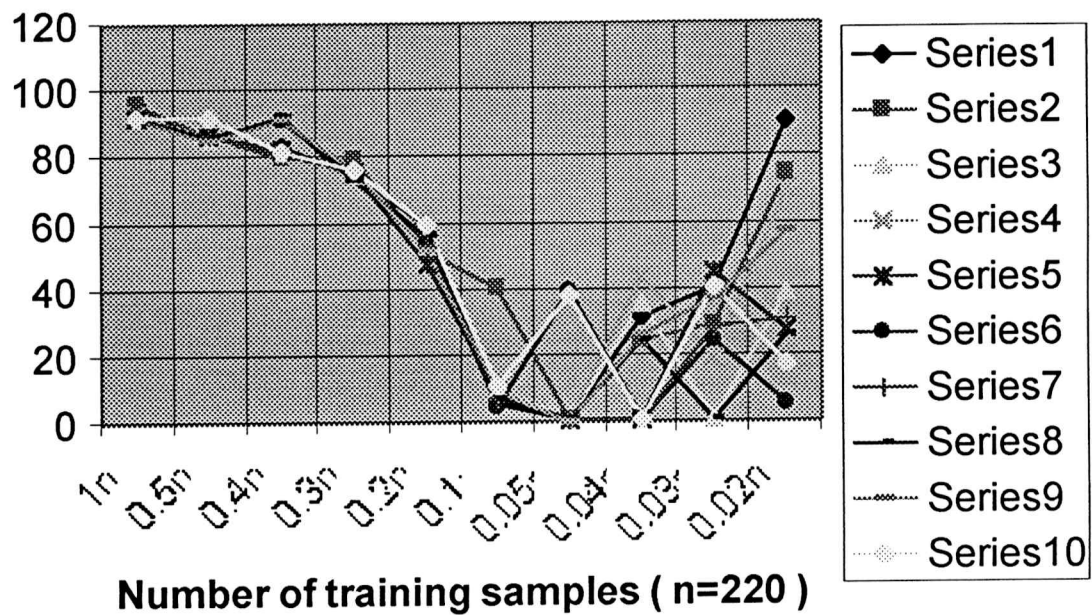
3) SoybeansNS.

This class has 364 number of samples. This is approximately 1.5n.

The following table shows the experimental results:

Number of samples		1n	0.5n	0.4n	0.3n	0.2n	0.1n	0.05n	0.04n	0.03n	0.02n
trial	1	96.2	86	83.2	74.7	59.9	5.2	1.1	31.9	40.7	90.4
trial	2	96.2	86.3	79.9	79.4	51.9	40.7	0.5	24.5	29.7	74.9
trial	3	92	86	81.6	75	55.8	9.9	0.5	35.7	1.1	38.8
trial	4	93.1	86.3	82.1	74.5	52.2	8.8	1.1	43.4	19	69.8
trial	5	92	86.3	82.7	75.8	47.3	8.5	0.5	0.8	45.1	27.5
trial	6	92	86.8	91.2	74.7	53.8	4.7	39.6	0.8	24.5	5.2
trial	7	91.8	86.3	81.3	76.1	53.3	7.1	0.5	0.8	28.8	30.8
trial	8	91.8	85.7	92	73.4	56.3	6.9	0.5	25	0.5	26.4
trial	9	92.3	85.2	91.5	76.4	51.6	9.3	0.3	26.4	37.6	56.9
trial	10	92	91.8	82.1	76.4	59.6	11.3	38.5	0.5	40.7	17.6

**Plot of training accuracy for
SoybeansNS against the number of
training samples**



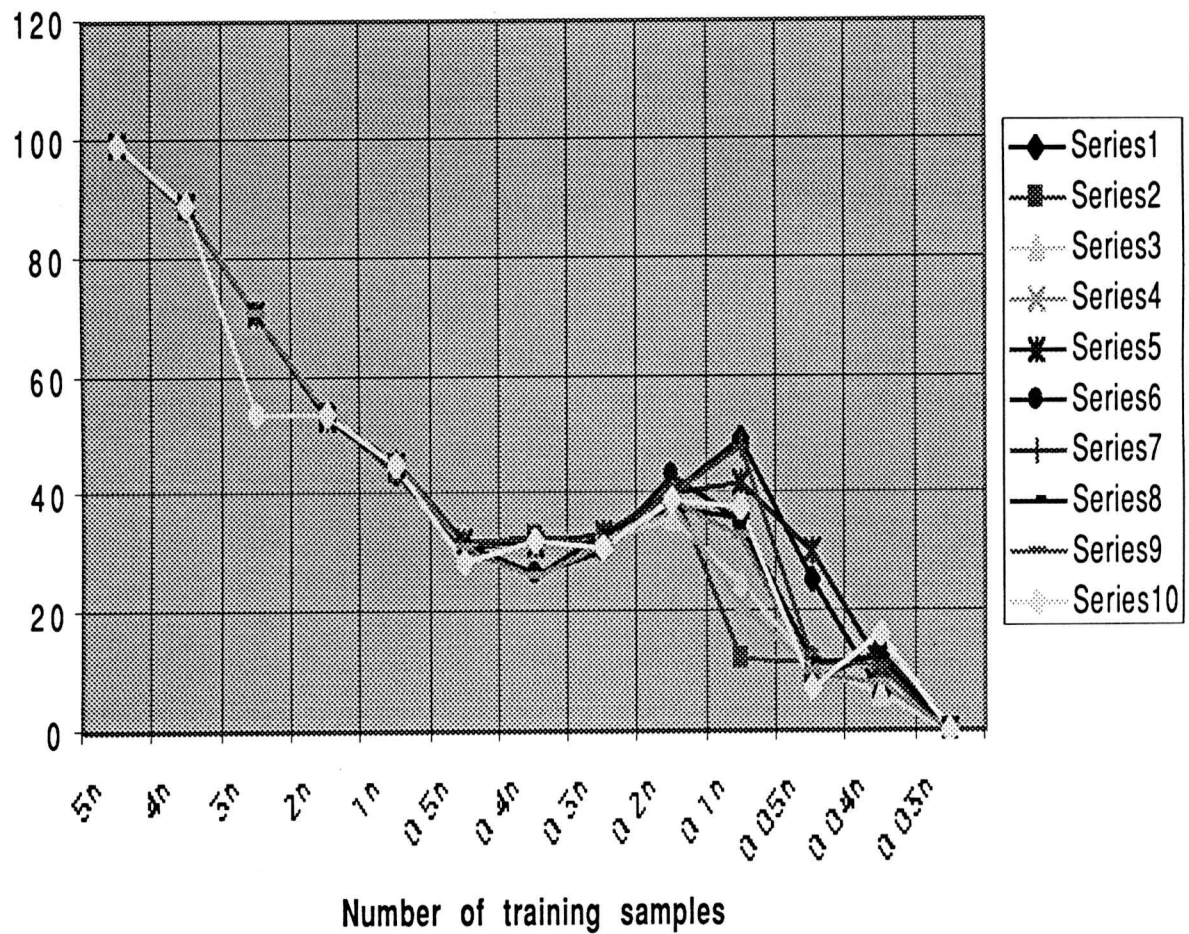
4) Corn cleantill EW.

This class has 1090 samples. This is approximately 5n.

The following table shows the experimental results:

Number of samples		5n	4n	3n	2n	1n	0.5n	0.4n	0.3n	0.2n	0.1n	0.05n	0.04n	0.03n
trial	1	99.1	89.2	70.6	53.3	43.7	31.1	32.1	31.8	40	49.4	25.5	6.3	0
trial	2	99.1	88.8	70.7	53.6	43.6	31.5	32.8	32.6	41.2	12.2	11.7	10.6	0
trial	3	99.1	89.3	70.6	53.2	45.2	32.6	31.4	32.1	35.7	25	10.7	6.3	0
trial	4	99.1	89	71	53.1	45.2	32.4	30.9	31.3	38.2	21.6	10.6	11.1	0
trial	5	99.1	89.1	70.6	52.8	45	32.3	31.9	33.5	39.9	42.2	30.6	11.2	0
trial	6	99.1	88.6	70.8	53.8	45	31.7	26.8	30.2	43.4	35.2	8.6	13.3	0
trial	7	99.1	88.6	70.5	52.9	45.5	32.2	32.6	31.7	40.2	48.2	12.4	11.2	0
trial	8	99.1	89.2	70.9	53.1	45	31.7	26.4	33.5	37.6	36.1	11.6	12.1	0
trial	9	99.1	89.1	70.9	53.7	45.5	31.3	26	31.9	38.9	32.7	9.5	7.9	0
trial	10	99.1	89.1	53.9	53.5	45	28.5	32.4	30.8	38.9	37.8	7.2	16.3	0

Plot of training accuracy for CornleantillEW against the number of training samples



5)Wheat.

This class has 1191 number of samples. This is approximately 5.5n.

The following table shows the experimental results:

Number of samples	5.5n	5n	4n	3n	2n	1n	0.5n	0.4n	0.3n	0.2n	0.1n	0.05n	0.04n	0.03n	0.02n
trial 1	100	97.1	97.1	97	97	96.1	92.2	93	93.7	94.1	95.1	96	94.8	77.5	85.7
trial 2	100	97.1	97.1	97	97	96.1	92.2	93.5	94.4	94.7	95	95	95.5	94.4	74.9
trial 3	100	97.1	97.1	96.9	97	95.3	92.2	93.5	94	94.2	95.4	92.2	94.4	86.9	38.8
trial 4	100	97.1	97.1	96.9	97	95.4	92.4	93.3	93.5	94.5	95.1	66.4	95	66.8	7.9
trial 5	100	97.1	97.1	97	97	96	91.7	93.5	93.7	94.1	95.5	95.6	95.6	66.1	3.3
trial 6	100	97.1	97.1	97	97	95.4	92.1	95	94	94.6	95.2	95	62	83	30.1
trial 7	100	97.1	97.1	96.9	97	95.5	92.3	93.3	94	94.1	95.5	91.4	95.6	93.8	28
trial 8	100	97.1	97.1	97	97	95.3	92.3	95.3	93.9	94.3	94.8	94.7	88.2	93	3.4
trial 9	100	97.1	97.1	96.9	97	95.4	95	95.7	94	94	95.2	88.4	61.8	54.2	87.8
trial 10	100	97.1	97.1	97	97	95.3	92	93	93.9	94.2	95.5	96.1	90.2	67	87.5

Figure 1 is a line graph showing the relationship between the number of training samples (X-axis) and the number of hidden nodes (Y-axis) for ten different series (Series1 to Series10). The X-axis values are 5⁵, 5⁴, 4⁵, 3⁵, 2⁵, 1⁵, 0.5⁵, 0.3⁵, 0.1⁵, 0.05⁵, 0.01⁵, 0.005⁵, 0.002⁵, and 0.001⁵. The Y-axis ranges from 0 to 120. The graph shows that as the number of training samples decreases, the number of hidden nodes generally decreases for most series, with some series (like Series1 and Series2) showing a sharp drop at very low sample sizes.

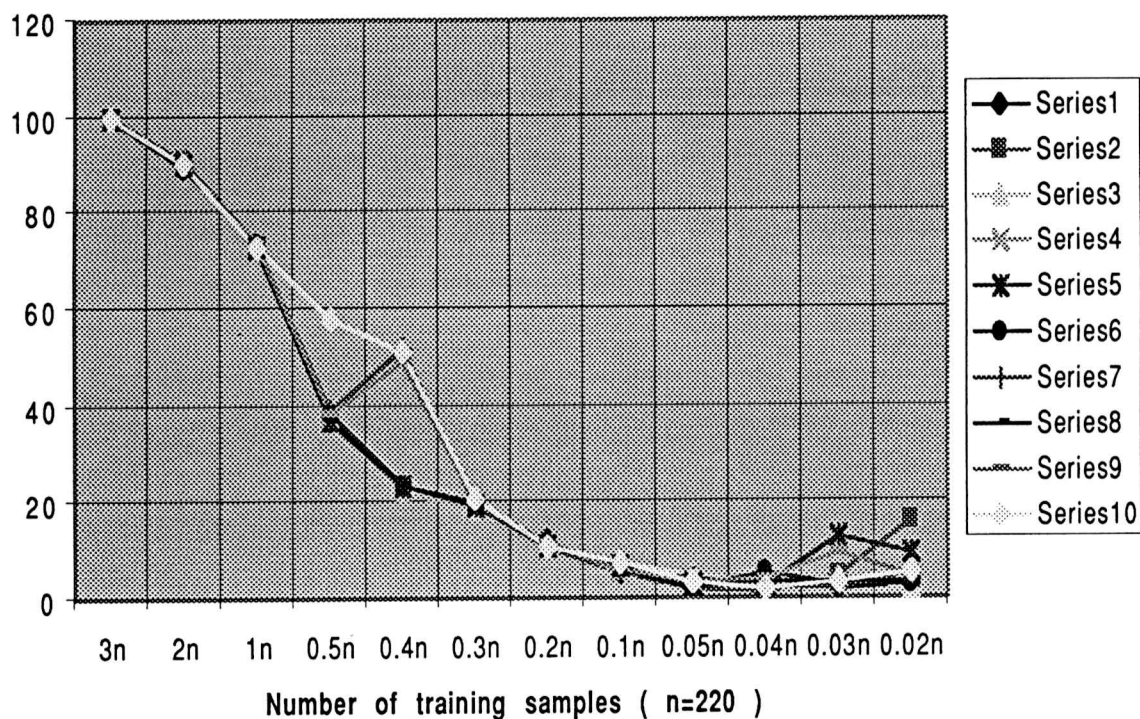
6) SoybeansmintillNS.

This class has 684 number of samples. This is approximately $3n$.

The following table shows the experimental results:

Number of samples		$3n$	$2n$	$1n$	$0.5n$	$0.4n$	$0.3n$	$0.2n$	$0.1n$	$0.05n$	$0.04n$	$0.03n$	$0.02n$
trial	1	100	89.9	73.2	37.3	23.1	18.9	11.8	5.6	3.9	4.2	2.3	2.9
trial	2	100	91.1	73	37.1	23	19.6	11.1	5.4	2.3	1.6	5	16.1
trial	3	100	91.1	72.7	37.6	23	18.9	10.8	7.5	4.4	3.5	3.9	2.2
trial	4	100	90.2	71.8	37.9	22.2	18.6	12.1	5.6	3.7	4.4	2.8	5.7
trial	5	100	90.4	73.1	37.4	23.5	19.4	11.1	7	2.6	2.5	12.6	9.5
trial	6	100	90.6	73	38.5	23.5	19.7	11.5	6.4	1.9	5.3	2.5	2.8
trial	7	100	90.8	73.2	39	51.9	19.3	11.3	5.8	3.8	2.5	2.2	5.4
trial	8	100	90.2	73.4	37	23.8	20.3	11.3	6.3	3.5	3.4	3.4	3.4
trial	9	100	90.6	73.7	38.3	49.3	20	11.7	6	3.7	3.8	9.6	4.4
trial	10	100	90.5	73.1	58.2	50.9	20.6	10.8	7.2	3.7	2.2	3.7	5.4

Plot of training accuracy for soybeanmintilINS
against the number of training samples



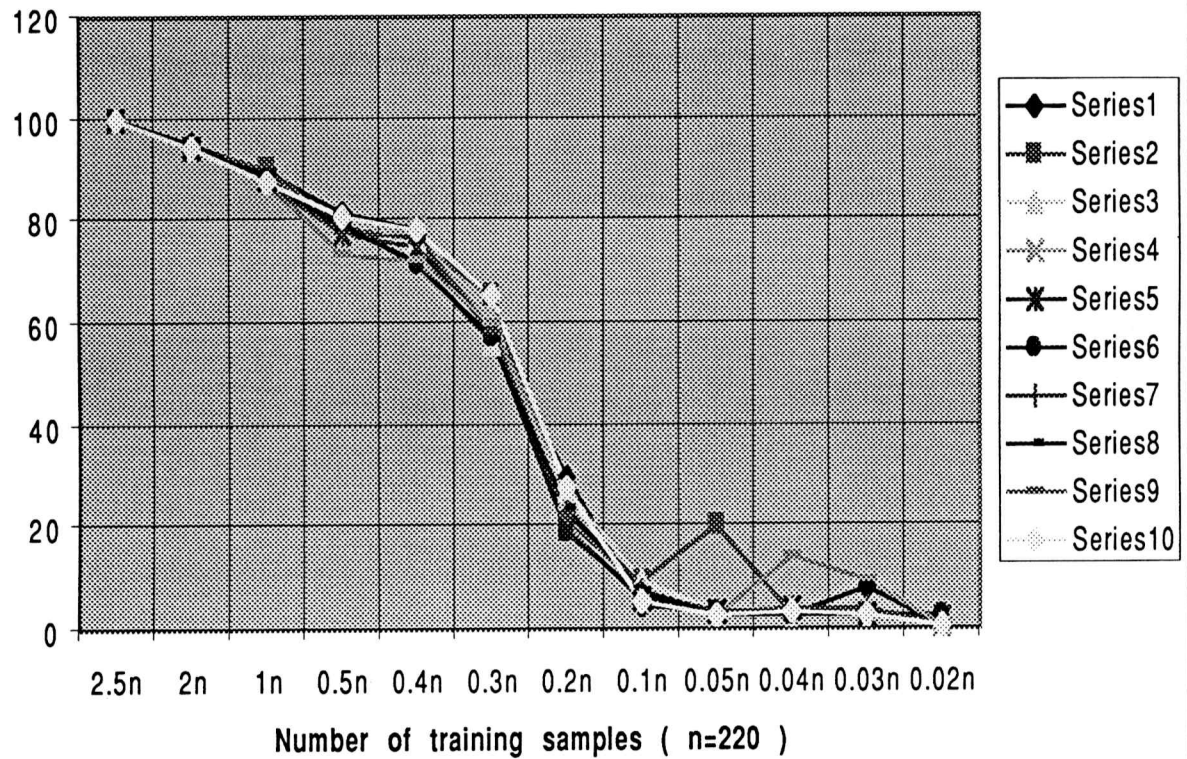
7) Soybeansdrilled.

This class has 585 number of training samples. This is approximately $2.5n$.

The following table shows the experimental results:

Number of samples		$2.5n$	$2n$	$1n$	$0.5n$	$0.4n$	$0.3n$	$0.2n$	$0.1n$	$0.05n$	$0.04n$	$0.03n$	$0.02n$
trial	1	100	95	89.9	81.7	78.3	58.1	29.6	4.8	3.1	2.6	8.2	0.7
trial	2	100	94.4	90.6	79	76.2	65.5	20.7	9.4	20.2	2.6	2.6	0.7
trial	3	100	94.5	88	77.4	76.4	56.2	26.7	9.4	3.1	2.6	2.4	1.2
trial	4	100	94.5	88.5	75.6	78.6	59.3	25.8	5.6	2.6	2.6	2.6	1
trial	5	100	94.7	88	77.4	74.5	58.8	19.3	5.8	3.1	4.3	2.6	2.1
trial	6	100	94.5	88	79.5	71.3	57.6	18.8	6.7	3.6	2.9	2.6	2.4
trial	7	100	94.7	87.5	77.6	77.1	58.5	22.6	7.4	3.4	4.3	4.1	1
trial	8	100	94.9	88.2	79.1	72	56.9	23.6	7.2	2.9	2.6	2.6	0.7
trial	9	100	94.7	87.9	73.8	71.8	59.8	28.2	5.6	2.9	14.4	9.6	0.7
trial	10	100	94.5	87.4	80.9	78.1	65.5	27.4	5.3	2.7	3.6	2.6	0.9

Plot of training accuracy for Soybeansdrilled
against the number of training samples



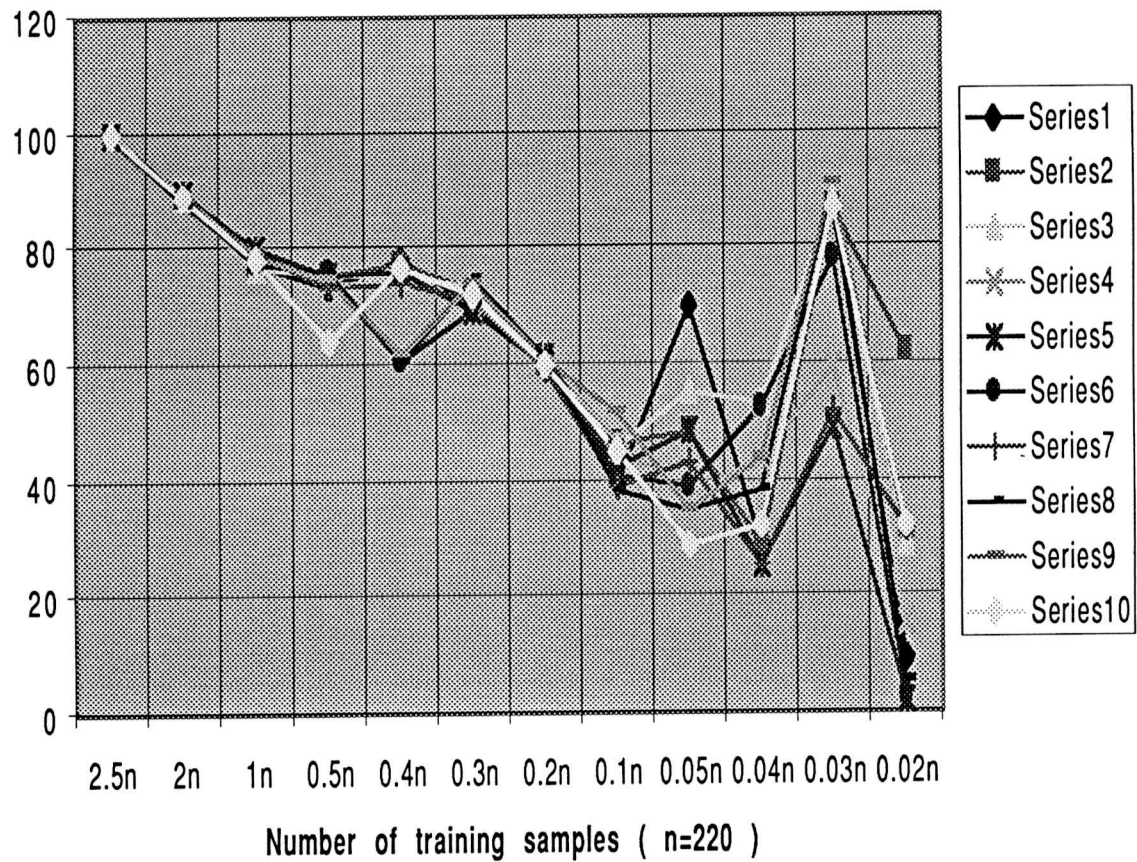
8) SoybeanscleantillEW.

This class has 567 number of training samples. This is approximately $2.5n$.

The following table shows the experimental results:

Number of samples		$2.5n$	$2n$	$1n$	$0.5n$	$0.4n$	$0.3n$	$0.2n$	$0.1n$	$0.05n$	$0.04n$	$0.03n$	$0.02n$
trial	1	100	89.2	77.4	73.5	78.1	70.9	62.3	39.5	70.4	28.9	86.8	9.2
trial	2	100	88.9	76.9	75.5	77.8	72.8	60.1	46.6	48.3	28.7	87.1	61.7
trial	3	100	89.1	77.1	75.8	75.7	73.07	62.3	45.9	55.4	53.3	88.2	28.7
trial	4	100	89.9	79.4	74.1	74.1	70.2	61.6	46.4	37.4	43.6	59.6	15.2
trial	5	100	89.8	80.2	75.3	76.5	69.3	61.2	42.3	48.3	25.4	49.2	2.1
trial	6	100	89.2	77.2	76.2	60.8	69.7	60.3	40.9	39.3	52.4	78	2.4
trial	7	100	89.2	79.4	73.4	73.9	71.3	59.6	39.2	42.9	25.4	51.7	31.9
trial	8	100	88.5	78.5	75.8	60.8	75.1	61.6	38.1	35.1	38.6	90.8	5.1
trial	9	100	89.4	79.2	75	62.1	73.9	60.7	51.7	35.1	43.7	90.8	30.3
trial	10	100	89.1	78.3	64.2	77.1	72	60	45	29.1	32.6	86.9	32.3

Plot of training accuracy for SoybeanscleantillEW
against the number of training samples



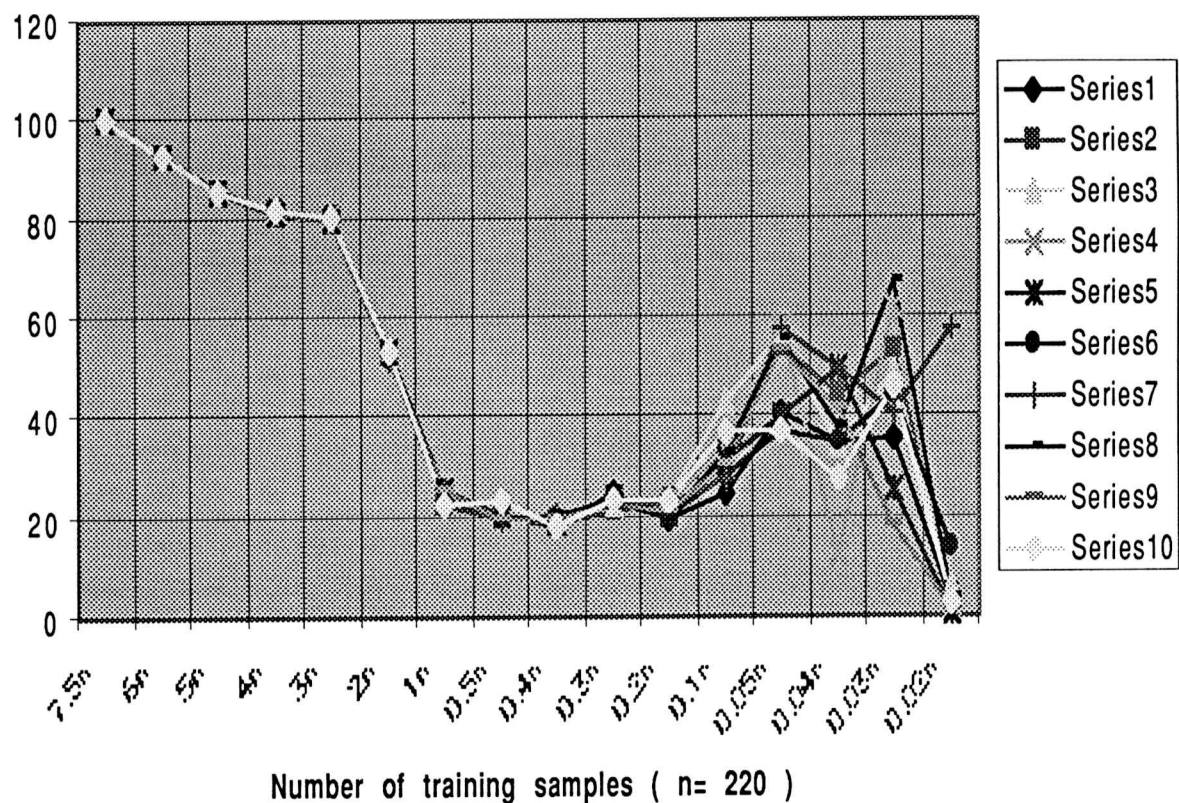
9) Corn.

This class has 1653 number of samples. This is approximately $7.5n$.

The following table shows the experimental results.

Number of samples	7.5n	6n	5n	4n	3n	2n	1n	0.5n	0.4n	0.3n	0.2n	0.1n	0.05n	0.04n	0.03n	0.02n
trial 1	100	93	85.9	81.6	80.3	53.7	22.7	20.4	18	24.9	19.8	24.7	41.1	34.8	35.5	2.9
trial 2	100	93	85.7	81.9	80.2	52.6	23	20.1	18.9	21.1	23.4	34.1	54.3	44.9	53.5	4.2
trial 3	100	93	85.9	81.7	80.7	53	25.8	19.4	18.8	22.4	23.5	43.4	56	34.2	49.1	6.2
trial 4	100	92.9	85.7	81.5	80.6	53.4	25.6	20.1	20	24.2	23.4	34.2	57.9	12.2	62	1.9
trial 5	100	92.9	85.9	81.5	79.7	53.2	25.7	20.3	18.9	24.1	20.4	31.8	40.6	49.6	25.8	0.8
trial 6	100	93.1	85.7	81.9	80.2	52.3	25.5	19.8	19.5	24.1	20.3	28.5	37.3	35.2	44.3	13.9
trial 7	100	92.9	85.9	81.4	79.8	52.1	25.8	20	20.3	23.6	22.7	27.9	57.6	49.6	40.8	57.5
trial 8	100	93	85.7	81.8	80.4	52.1	25.3	19.8	20.6	23.5	20.1	33	55.9	38.4	67.3	3.4
trial 9	100	92.9	86	81.5	79.9	51.6	25.5	20	19.7	22.8	21.5	30.5	53.5	41.4	18.4	1.3
trial 10	100	93	85.7	81.7	80.3	53.2	22.7	23.1	18.6	23.6	23.4	37.6	37.4	27.5	46.3	3

Plot of training accuracy for Corn against the number of training samples



10) CorncleantillNS.

This class has 2770 number of training samples. This is approximately

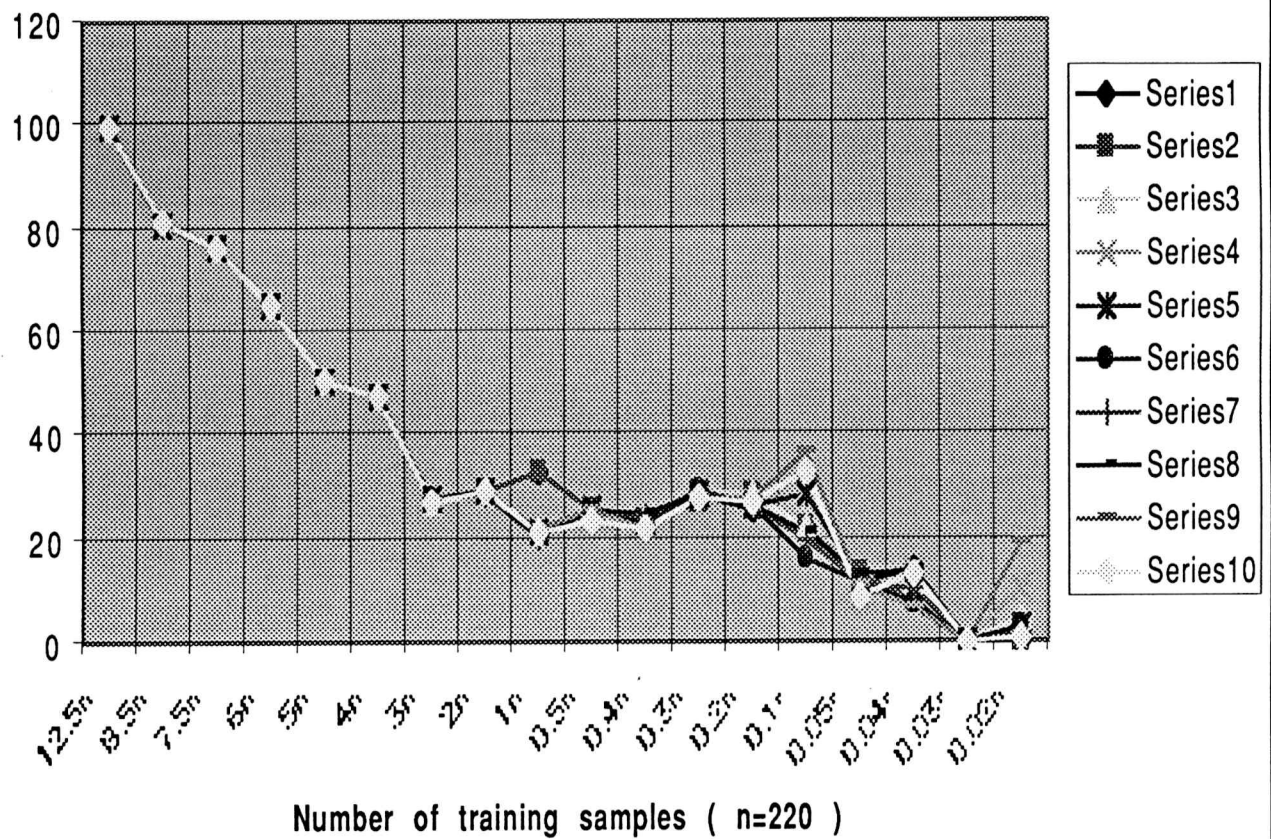
12.5n.

The following table shows the experimental results:

Number of samples		12.5n	8.5n	7.5n	6n	5n	4n	3n	2n	1n
trial	1	99.6	80.9	75.9	65.1	50.5	47.8	27.4	29.4	32.7
trial	2	99.6	80.8	75.9	65.1	50.5	47.8	27.4	29.4	32.8
trial	3	99.6	80.9	75.9	65.1	50.5	47.9	27.4	28.7	21.3
trial	4	99.6	80.8	75.9	65.1	50.5	47.8	27.7	29.1	21.1
trial	5	99.6	80.8	75.9	65.2	50.5	47.8	27.6	29.4	21.1
trial	6	99.6	80.8	75.9	65.2	50.4	47.8	27.5	29.2	21.4
trial	7	99.6	80.8	75.9	65.1	50.5	47.8	27.5	28.8	20.5
trial	8	99.6	80.8	75.9	65.2	50.4	47.8	27.5	29.1	20.8
trial	9	99.6	80.8	75.9	65.4	50.5	47.9	27.7	28.8	21.8
trial	10	99.6	80.9	75.9	65.2	50.5	47.8	27.4	29.2	21.2

Number of samples		0.5n	0.4n	0.3n	0.2n	0.1n	0.05n	0.04n	0.03n	0.02n
trial	1	25.8	23	28.6	25.5	22.2	11.7	14.2	0	0.2
trial	2	25.3	23.5	26.9	28.6	22.3	13.1	8.1	0	0.3
trial	3	26.2	24.2	29.1	28.4	21.6	10.1	11.4	0	4.1
trial	4	26.6	22.6	29.5	27.4	29.6	8.7	13.8	0	10.4
trial	5	25.4	23.3	28.8	26.4	28.8	11.6	9.9	0	2.6
trial	6	25.7	22.4	29.6	26.6	16	11.7	7.1	0	2.6
trial	7	25.3	23.8	29.2	26.7	19.5	11.7	9.9	0	1.5
trial	8	25.4	24.9	28.8	26.6	21.3	12.9	13.6	0	0.2
trial	9	25.8	21.8	28.1	27.6	36.8	10.9	6.3	0	19.1
trial	10	23.6	22.3	27.7	26.8	33.1	8.9	13.1	0	0.8

Plot of training accuracy for CorncleantilINS
against the number of training samples



Discussion:

Remember the 2 questions we are interested in? First is the question of the validity of classifying a given class of interest with as few as 3 samples using the Leave-One-Out statistics.

The other one is whether we could come out tentatively with a minimum number of samples in relation to the original number of features or dimensions so as to be able to give a relatively fair amount of classification accuracy. Of course, the definition of “relatively fair” is a context that differs from one person to the other. However, we are going to look at the plotted results and the trend derived from those graphs and hopefully we could answer some of these questions.

Let us define “divergence point” as the point where the training accuracies diverges in the plot.

First, let us look at the “divergence point” of these plotted results:

- 1) Building diverges after $0.1n$
- 2) CornmintillEW diverges after $0.1n$.
- 3) SoybeansNS diverges at $0.1n$
- 4) CorncleantillEW diverges at $0.1n$.
- 5) Wheat diverges after $0.1n$
- 6) SoybeansmintillNS diverges at $0.04n$.
- 7) Soybeansdrilled after $0.1n$
- 8) SoybeanscleantillEW diverges at $0.1n$

Up to this point, we see kind of a trend! That is the “divergence point” of the plots of training accuracies against the number of training samples seems to be near or at $0.1n$ (i.e. 22 samples).

Next look at what is the accuracy when the number of training samples are $0.02n$ (ie. approximately 4 number of samples). This should roughly answer one of our questions. That is the validity of using as few as 3 samples to classify a given class of interest.

Skimming through all the plots, the accuracies vary from extremely high to extremely low in some cases, almost zero in some cases and very low accuracies in some cases.

Hence, even though we might sometimes get very high accuracies using very limited number of samples, the results could vary erratically and thus to be on a safe side, choose as many training samples to do the classification. However, from the experiment that I have conducted, using at least more than $0.1n$ (22 samples) would at least prevent us from reaching the “divergence point”.

Appendix A

soybeanscleantillEW
74-5

corncleantillNS
75-2

soybeansmintillNS
75-3??

soybeansdrilled
75-4

corncleantillEW
74-6

corncleantillNS
75-9

corn
75-5

cornmintillEW
74-14

corncleantillNS
75-10

corncleantillNS
75-11

corn
75-6

wheatbuilding
74-135-19

corn
75-7

wheat
74-22

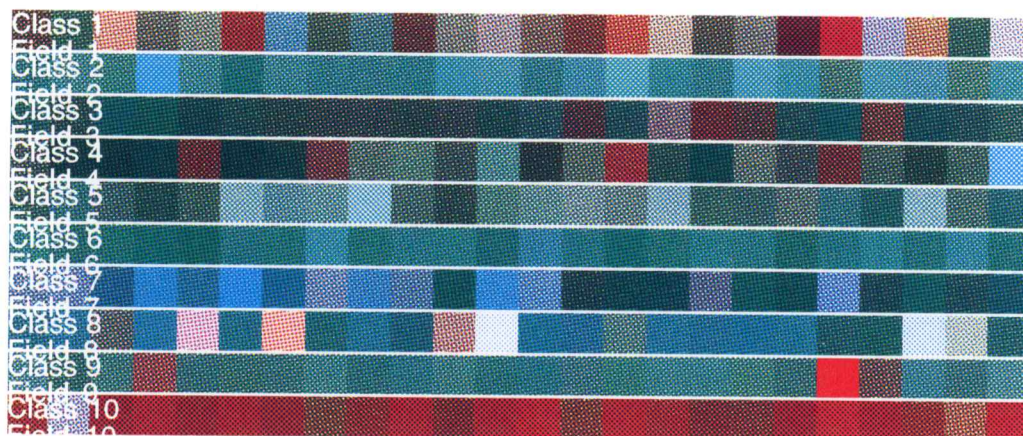
soybeansNS
75-14 (p1)

corncleantillNS
75-15

corncleantillEW
75-16

soybeansNS
75-14 (p2)

corn
75-17



Appendix B

The next page contains the results of training accuracies for the 10 classes using 100% of all the pixels for each class

TRAINING CLASS PERFORMANCE (Resubstitution Method)

Project Class Name	Class Number	Reference Accuracy+ (%)	Number Samples	Number of Samples in Thematic Image Class		
				0 background	1 building	2 corncleantil cornclea
building	2	92.0	112	0	103	0
corncleantilleW	3	100.0	1090	0	0	1090
corncleantillNS	4	99.6	2770	0	1	4
cornmintilleW	5	100.0	220	0	0	0
corn	6	99.8	1653	0	0	0
soybeanscleantilleW	7	100.0	567	0	0	0
soybeansdrilled	8	100.0	585	0	0	0
soybeansmintillNS	9	100.0	684	0	0	0
soybeansNS	10	100.0	364	0	0	0
wheat	11	100.0	1191	0	0	0
TOTAL			9236	0	104	1094
Reliability Accuracy (%)*					99.0	99.6

OVERALL CLASS PERFORMANCE (9212 / 9236) = 99.7%
Kappa Statistic (X100) = 99.7%. Kappa Variance = 0.000000.

*The rest of Appendix B is not
included in this pdf file*