Target Confirmation in LWIR Hyperspectral Data using Neural Networks

Owen Thomas
Thales Optronics
Thorpe Road, Staines, Middlesex. TW18 3HP

Abstract

Neural networks are applied to the classification of targets in LWIR hyperspectral data gathered with the AHI sensor. Of primary interest is a feed forward architecture trained in a single iteration through the use of a virtual layer and alternating direction singular value decompositions. Training that may take hours for MLPs with particular training sets may require only seconds for the single iteration process.

Keywords: Alternating direction singular value decomposition, multilayer perceptron

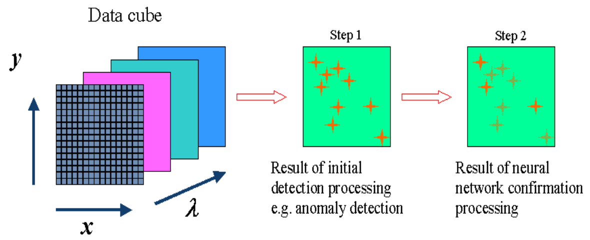
Introduction

The identification of targets using spatial information is not possible at extreme range when the target occupies a single pixel or less and can be extremely difficult in regions of high spatial clutter. Analysis of the pixel radiation in tens or even hundreds of narrow bands can reveal unique features that are characteristic of a particular object type.

Neural networks are a proven technology in pattern classification and can be effective in difficult scenarios. The networks would be trained 'off-line' with representative data for the anticipated mission.

Discussion

The neural network target confirmation processing is envisaged as forming the second stage of a two-stage process. The first stage filters the data to eliminate most of the background with, for example, an anomaly detector. The targets and anomalous pixels should pass to the second stage, that stage classifies the pixels into target/non-target. The architecture ultimately envisaged is illustrated in the following:



Processing architecture

This paper considers neural network processing of entire scenes. The initial detection preprocessing stage is omitted but a dimensionality reduction exercise is reported. Given the nature of the objects in the scenes selected here the preprocessing might comprise canonical discriminant analysis, employed effectively by Thales Optronics on other projects.

Neural network architectures

The neural networks considered in this project are supervised. The required network output for a particular input class is specified.

The principal architectures in this group are the multilayer perceptron (MLP) and radial basis function networks. The MLP form is represented in the following figure.

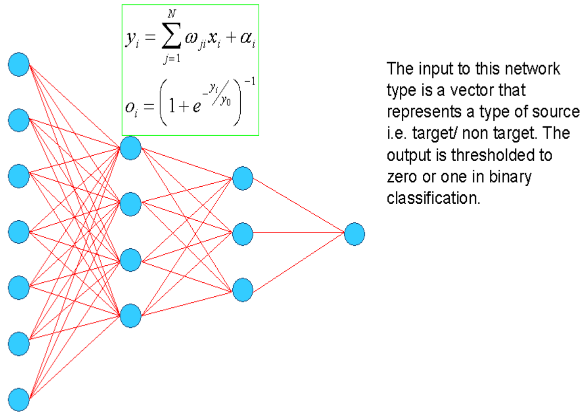


Illustration of the architecture of the multilayer perceptron with 2 hidden layers

Alternating singular value decompositions

The architecture of the ADSVD NN is illustrated below. There are input, hidden and output layers as with MLPs but there is an additional *virtual* layer between the input and hidden layers.

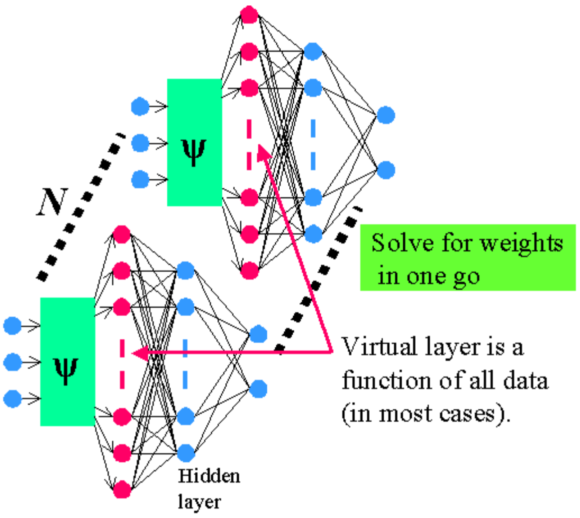


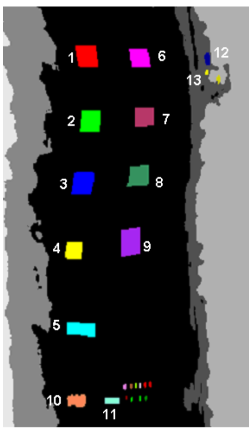
Illustration of the architecture of the ADSVD NN

The weight-matrix inverses required in the training are effected through application of SVD. The two such inverses in the scheme with a single hidden layer are from the right and then from the left – hence the term alternating direction singular value decomposition.

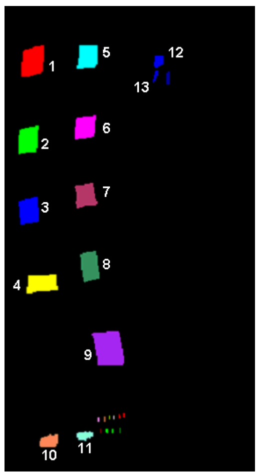
Application of the ADSVD NN to hyperspectral data analysis has been reported in the literature (Subramanian et al (1997)). That investigation employed sub-networks to identify individual classes and reject the others. In the present work single networks are trained to classify imagery.

The data analysed

The data analysed thus far on this project were gathered at a trial in a tropical location. The data were captured from an aircraft at around 1000 feet using the University of Hawaii AHI sensor operating in the LWIR region. The data is of 25 bands although the sensor is capable of gathering several hundred. The training data used in the experiments reported here were extracted from the scene outlined below.



Training scene



Testing scene

The black and grey areas represent background/non-target regions while the coloured pixels indicate targets whose locations are supplied in a file associated with the data cube (as is the colour scheme). The black area is open grassland/sand (a cleared region of rainforest). At the left and right edges are wooded areas and between the open ground and the right woods appears to be a track used by vehicles on the trial. The eleven largest coloured blocks in the black region and the blue and left yellow blobs towards to top right corner represent the 'targets' from which training spectra were extracted. Background training spectra were extracted from the different background types. A total of 2452 training spectra were employed. The test data comprises a look at the same scene a few days later. One of the panels in the training data is no longer present but has been replaced by a larger panel in a location just to the right.

Preprocessing

The "resolved" targets in the data sets are not best preprocessed using anomaly detectors. Canonical discriminant analysis (CDA) may be used to maximise the variance between classes while also minimising the within class variances. The required transformation is the solution of a generalized eigenvalue problem. Such a technique was employed by Subramanian et al in their work and has been employed by Thales Optronics Ltd with and without subsequent NN processing on other projects. In the present case a similar mathematical technique has been deployed to maximise the signal to noise ratio in the data. There is marked vertical striping in the output of numerous processes applied to these data. Consequently a noise covariance matrix is calculated from the hyperspectral cube in which each band is filtered with a vertical edge detector. A minimum noise fraction transformation (Barhen, Cogswell and Protopescu (2000)) is calculated as a solution to a generalized eigenvalue problem in which the second covariance matrix is calculated from the raw data. If Σ is the covariance matrix of the data and Σ_n is the covariance matrix of the filtered data then the minimum noise fractions are the inner products of the spectra with the eigenvector solutions of $\Sigma a_i = \lambda_i \Sigma_n a_i$.

Results

Neural networks, in the work reported here, have been employed to classify the pixels in a scene without a prefiltering process designed to reject background pixels while retaining targets e.g. CDA. An MLP and an ADSVD NN have been trained to classify the pixels in a scene. The outputs of the application of these have been analysed to determine how well the networks correctly *recognize* the target pixel type and also to determine how well the networks *detect* a target pixel correctly. The full data set is analysed and the results are compared to the analysis of minimum noise fractions. All spectra are normalized to be of unit length prior to presentation to the networks.

The first result shows how well an MLP with 6 hidden nodes has recognized the target type. The training time required is over four hours.

Label of training object

Label of test scene object	1	2	3	4	5	6	7	8	9	10	11	12	13
1	471	0	0	0	0	0	0	0	0	0	0	53	0
2	0	2	173	0	189	0	0	1	0	0	0	2	0
3	1	0	367	0	0	0	0	0	0	0	0	38	0
4	10	0	0	0	0	0	1	1	0	0	242	0	164
5	0	0	0	0	0	0	0	1	0	0	56	0	277
6	0	0	0	0	0	0	0	0	0	0	40	1	216
7	0	0	0	0	0	0	0	0	0	0	48	3	275
8	0	0	0	0	0	0	0	0	0	0	397	1	8
9	26	0	0	0	0	0	0	0	0	0	677	76	10
10	0	7	2	0	0	0	0	0	0	163	0	0	0
11	10	0	0	0	0	0	0	0	0	0	18	25	43
12	1	0	0	0	0	0	0	0	0	7	0	0	0
13	0	0	0	0	0	0	0	0	0	10	0	0	0

Part confusion matrix for application of MLP to 25 bands

Targets 1, 3 and 10 are identified fairly well.

The ADSVD NN results that follow are all produced with 150 nodes in the virtual layer and 75 nodes in the hidden layer – training time is around 16 seconds. The next table shows the labels attached to the objects in the testing scene.

Label of training object (columns); **Label of test scene object** (rows)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	50	0	0	0	0	0	0	0	0	0	0	474	0
2	0	2	11	1	289	3	6	22	0	0	0	0	0
3	0	0	406	0	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	221	3	0	169	1	0
5	0	0	0	0	0	0	0	304	0	0	0	1	0
6	0	0	0	0	0	0	0	299	0	0	16	2	0
7	0	0	0	0	0	0	0	210	1	0	110	34	0
8	0	0	0	0	0	0	0	1	4	0	401	0	0
9	35	0	0	0	0	0	0	0	0	0	468	286	0
10	0	0	0	0	0	0	0	0	0	172	0	0	0
11	1	0	0	0	0	0	0	0	0	0	3	90	1
12	0	0	0	2	0	0	0	0	0	4	0	0	0
13	0	0	0	0	0	0	0	0	0	10	0	0	0

Part confusion matrix for application of ADSVD NN to 25 bands

Using all 25 bands the ADSVD network has really only recognized targets 3 and 10.

The ADSVD network has been retrained using only the first five minimum noise fractions and the results for this experiment are seen below.

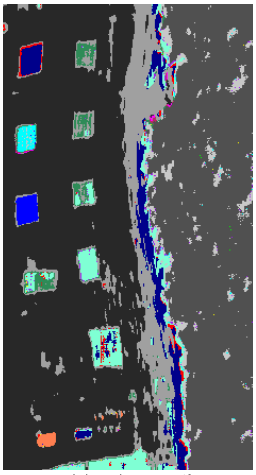
Label of training object (columns); **Label of test scene object** (rows)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	508	0	0	0	0	0	0	0	0	0	0	15	0
2	0	418	0	0	0	0	0	0	0	0	0	0	0
3	0	0	406	0	0	0	0	0	0	0	0	0	0
4	246	0	0	80	5	0	0	0	14	0	80	3	0
5	0	0	0	1	0	0	15	3	2	0	253	45	0
6	3	0	0	12	0	0	28	35	24	0	229	13	0
7	5	0	0	1	0	0	9	11	4	0	304	21	0
8	7	0	0	2	6	0	0	0	175	0	203	13	0
9	247	0	0	0	0	0	0	0	0	0	428	113	0
10	0	0	0	0	0	0	0	0	0	172	0	0	0
11	0	0	0	0	0	0	0	0	0	0	4	91	0
12	1	0	0	0	0	0	0	0	0	7	0	0	0
13	0	0	0	0	0	0	0	0	0	9	0	0	0

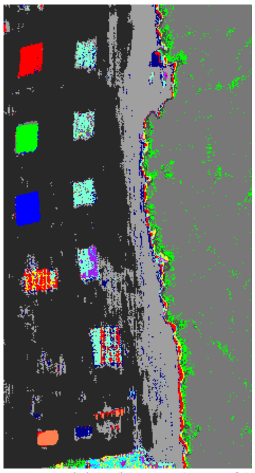
Part confusion matrix for application of ADSVD NN to 5 mnfs

This network produces the best results. Targets 1,2 3 and 10 are easily recognized while target 4 and target 8 (labelled target 9 in the training set) are partially recognised.

Recognition images for the ADSVD networks are seen next.



Recognition image (25 bands)



Recognition image (5 mnfs)

Performance figures for target detection using the same networks have been generated. A target pixel is regarded as detected if its declared class is that of any target. Detection is good as can be seen from the recognition images – the targets are coloured while most of the background is grey. In the table Pd is calculated as number of true detections divided by number of target pixels (including those excluded from the training set) while Pfa is number of non-target pixels declared as target divided by total number of background pixels.

	Pd	Pfa
MLP 25 bands	0.9403	0.1003
ADSVD 25 bands	0.9394	0.0886
ADSVD 5 mnfs	0.9710	0.0776

Detection performance of the networks

Conclusions

The ADSVD NN has been shown to produce results comparable or, with modest preprocessing, even superior to those of MLPs. The great advantage of this network is that the training times required are of the order of seconds rather than hours for the MLPs.

The twenty-five bands seen in this investigation are few in number for a hyperspectral data set. Processing performance with more bands and fewer artefacts would be expected to be rather better.

Acknowledgements

Thales Optronics is grateful to DSTL and Ron Ballingall in particular for supplying the data employed on this project.

References

[1] Subramanian et al, 1997, SPIE 3071, 128-137

[2] Barhen J., Cogswell R. and Protopescu V., 2000, Neural Processing Letters 11, 113-129

[3] Hilger K., Stegmann, M. and Larsen R., 2002, Proc. MICCAI 2002, 444-451