Target Confirmation in LWIR Hyperspectral Data using Neural Networks.

O. Thomas
Thales Optronics
Staines, TW18 3HP

Abstract

Neural networks are applied to the classification of targets in LWIR hyperspectral data gathered with the AHI sensor. The networks may be trained extremely quickly in a single iteration. Performance is compared to that of an unmixing approach.

Keywords: Singular value decomposition

Introduction

This paper reports the work undertaken on an extension to the programme that was reported upon last year [1].

The networks of the first investigation were known as alternating direction singular value decomposition (ADSVD) neural networks. This was down to that the requirement in the training process to invert non-square matrices using SVD from the right and from the left. In between programmes an alternative to ADSVD was encountered. The alternative used here known as *DeepNet*, by essentially the same authors, is approximately equivalent to ADSVD without the hidden layer. A network that dispenses with this layer provides a truly ultrafast neural network. This extension employs solely this simpler network.

The data employed in this extension were captured using the same sensor as that that gathered the data for the initial investigation but the higher spectral resolution mode on this occasion has succeeded in delivering quality outputs and so this work has not needed to apply any preprocessing for noise reduction. The data are nominally of 256 bands and of rows of 256 pixels but there are "bad" pixels together at the end of the rows and some of the bands contain anomalous data.

The first investigation concentrated on a single scene comprising sheets and nets of many types arranged on a grid on open ground. The altitude from which the data were captured was around 1000 feet. The present investigation has examined similar scenes but also a scene containing military vehicles.

Against the scenes examined here a spectral unmixing approach has also been applied to provide a comparison between the performance achieved by the networks and that of a more conventional method.

DeepNet

The pursuit of ever-faster algorithms has motivated the application of a version of ADSVD in which there is no hidden layer [2,3]. The training requires only singular value decomposition from the left. A prior examination of this simpler approach has shown that the loss of the hidden layer has a negligible impact on classification performance. This algorithm has been implemented by the developers in a computer code known as *DeepNet*. A description of the algorithm follows.

Suppose the number of bands is I and that the number of output classes is O. The number of training patterns K should typically be rather larger than I. The patterns are stored in a matrix Ω_{KI} while the output labels are stored in a matrix R_{KO}. Two successive non-linear transformations map Ω_{KI} into the $K \times V$ matrix H_{KV} output by the virtual layer where V is the number of nodes in the virtual layer.

The published description of DeepNet specifies that V should be equal to K so that H_{KV} becomes a non-singular square matrix. The single weight matrix required to be generated in the training process is then

$$W_{VO} = inv\ (\varphi\ (\psi\ (\Omega))) \times \varphi^{-1}(R_{KO})$$

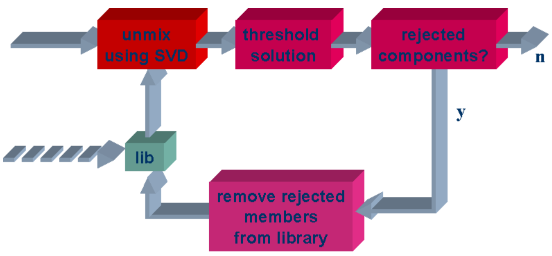
φ is the sigmoid function and ψ is the mapping that defines the virtual layer.

The matrix inversion may be implemented using singular value decomposition from the left.

In practice V need not be equal to K and this is exploited in the generation of the results seen later.

Unmixing approach

The unmixing approach, known as iterated singular value decomposition, is described in the following diagram.



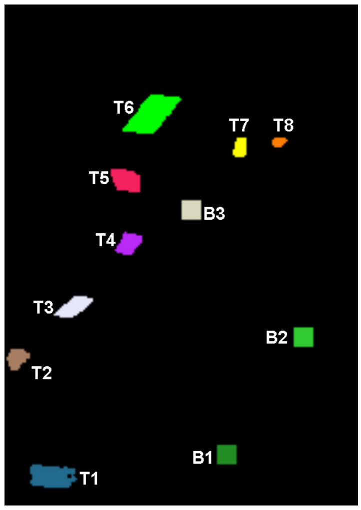
Outline of unmixing algorithm

In this unmixing approach the mixing equation is solved repeatedly and each time spectra are removed from the library matrix that correspond to an unmixing fraction less than a specified threshold.

The process ends when the unmixing fractions all exceed the threshold. There is no constraint to sum to unity.

The data analysed

The first application was designed to examine a grid of materials at a relatively low altitude. The training cube was captured at 1300 feet while the testing cube was captured at 1500 feet.

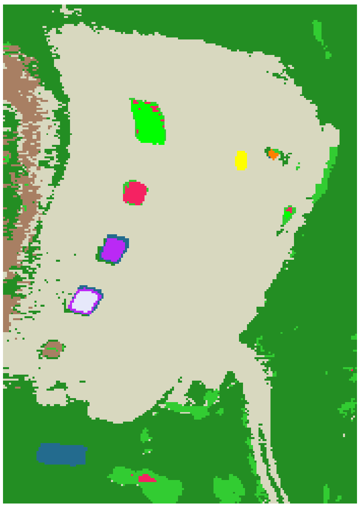


First training scene

The items/regions from which training spectra have been extracted are indicated in this representation of the training scene. Regions T1 to T8 represent objects in the scene (there are additional smaller items present) while regions B1 to B3 represent background types. The sensor here moves from the top of the scene to the bottom.

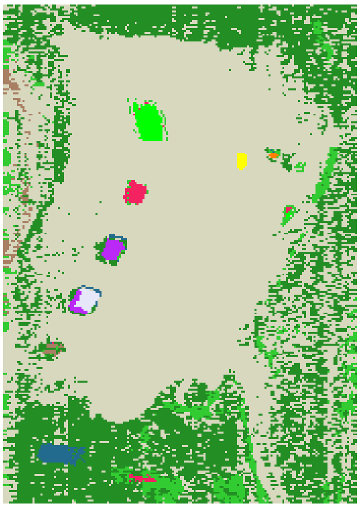
The testing phase has applied both DeepNet and iterated singular value decomposition (ISVD). The algorithms are trained using data from the training cube described above and tested on a second cube captured five days later. The NN is trained using 1422 normalized spectra, 500 virtual nodes and no hidden modes. The ISVD algorithm uses a library comprising 11 unnormalised spectra and a threshold of 0.2. Pixels analysed using ISVD are classified as the class exhibiting the greatest unmixing fraction. In total, 230 of the potential 256 hyperspectral bands have been used. The pixel values in the 26 discarded bands demonstrate odd behaviour and are considered unreliable.

The classification image for the test cube is shown below. The distortion in this test cube is not as severe as that in the training cube. The five panels towards the left on the gravel area form a clear row and almost all the pixels on these are correctly identified. In fact all the T_i are broadly identified correctly.



NN classification image

The unmixing result is "grainier" than the neural network result and is seen below.

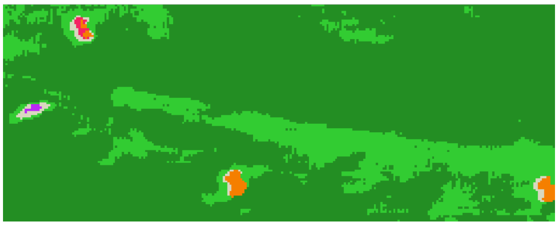


Spectral unmixing classification image

In the next experiment the training and test cubes were gathered on the same day around 4 hours apart and from the same altitude of 3000 feet. In the training scene below, vehicles V3 and V4 are of the same type. A dirt track runs left to right between V1and V2 and above V3 and V4. The track crosses open grass and at the bottom of the scene there is forest. V2, V3 and V4 have their engines running in the training cube. All engines are off in the test cube. The sensor in this case moves from left to right.

Vehicles in second training scene

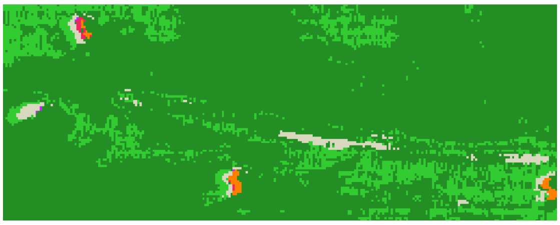
There are three target types and three background classes identified giving a total of six classes. In the case of the identical vehicles training pixels are extracted from both. In all 298 training pixels are employed exceeding the number of bands used by a little under a third. 250 virtual nodes have been employed and there is no hidden layer.



NN classification image

The grass is represented by the light green shade, forest by the dark green and the wheat colour represents dirt track. The background has been largely classified as such but of the wrong type and the track through the centre of the scene has been missed (classed as grass). The military vehicles have been fairly well identified. In the case of V1 the labelling is a mixture of true class and that of V3. The classifications of V2, V3 and V4 are purer but in the case of V2 it seems some on target pixels have been missed. In the immediate vicinity of the vehicles the churned ground (bare earth).

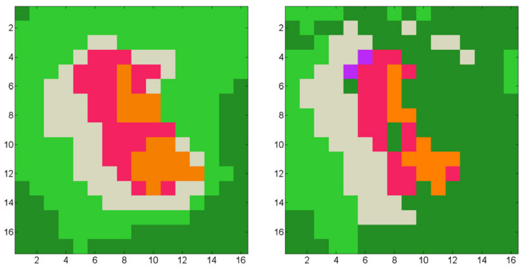
As in the analysis of the first scene ISVD has also been applied to provide a basis for comparison of the performance of the network. A smaller subset of signatures has been used to derive a library of six signatures and a threshold of 0.2 has been employed. The ISVD classification image is shown below.



Unmixing classification image

The performance against the targets is mixed and on the whole not as effective as that of the neural network. Pixels on the dirt track through the centre of the image have been identified and this could be crucial in some applications. Perhaps surprisingly the pattern of misclassification of the open grass as forest in the upper portion of the scene is similar to that of the neural network.

The classification of vehicle V1 is examined more closely in the figure below.



Comparison of classifications of V1

Conclusions

This paper has described an investigation to apply rapidly trained neural networks to the task of identifying targets in LWIR hyperspectral scenes. The performance of the neural network is found to exceed that of the unmixing approach used but the differences are relatively minor.

The data used in producing the results described here were captured from altitudes in excess of those from which previously seen data were measured. However, ensuring that the training data in the experiments are captured from the same or similar attitudes to those of the testing data has mitigated the effects of increased atmospheric attenuation and path radiation.

Some of the more difficult pixel contents to discriminate have been the natural vegetations. Grass and tree canopies are often confused - with grass commonly labelled as tree

References

1 Thomas, 0., 2004, Proc. EMRS DTC 1st Annual Technical Conference, **C19**

2 Barhen J. and Protopopescu, V., 2000, Proceedings of the 5th International Symposium on Distributed Robotic Systems, October 4th - 6th, Knoxville, Tennessee, p. 403

3 Barhen J., 2001, Proceedings of ESS'2001 European Simulation Symposium, October 18th - 20th, Marseille, France

Acknowledgements

The work reported in this paper was funded by the Electro-Magnetic Remote Sensing (EMRS) Defence Technology Centre, established by the UK Ministry of Defence and run by a consortium of SELEX Sensors and Airborne Systems, Thales Defence, Roke Manor Research and Filtronic.

Thales Optronics is grateful to DSTL for supplying the data employed on this project.