Optimizing the Non-linear Gamm ay Domain i Gamma-ray Astronomy using a Ne l Network Classifier

A.A. Chilingarian[1] and M.F. Cawley[2]

[1] Yerevan Physics Institute, Republic of Armenia
[2] St. Patrick's College, Maynooth, County Kildare, Ireland

ABSTRACT: We describe a neural network classifier which is used to find the optimal non-linear gamma-ray domain in multidimensional parameter space. A visualization approach based on the study of all possible 2-dimensional projections of the gamma-ray domain is used to delineate the boundary of the domain and explain the physical significance of the shape.

1 INTRODUCTION

The most sensitive detections of VHE gamma-ray sources to date have been made by the Whipple Collaboration using their "supercuts" analysis technique on data obtained using a 109-pixel Cherenkov imaging camera (Punch *et al.* 1991). This analysis approach uses a relatively simple linear region in 4-dimensional parameter space to define the gamma-ray domain (see section 2 for details). It is most unlikely that such a simple shape is optimal for signal enhancement. Numerous alternative analysis approaches have been tried on the Whipple Collaboration data (cluster analysis, singular value decomposition, neural network analysis, etc. — see review in Fegan *et al.* 1994) but the simple "supercuts" approach has proved surprisingly robust. In this paper we present an alternative approach to applying neural network classifiers to this type of data, and show that some improvement in sensitivity is obtainable by departing from the hypercubic shape. A simple method is used for visualizing and interpreting the output of the neural network.

2 THE NEURAL NETWORK CLASSIFIER

The conventional approach in applying neural network classifiers to Cherenkov imaging data is to first train the network using two classes of know events — usually simulated gamma-ray images and either simulated nucleonic images or real off-source images (Chilingarian, 1991; Reynolds, 1991; Hillas and West, 1991). The optimal network parameters are selected on the basis of maximising a "classification score", ie. the ability of the network to assign each type of event to the correct class. The main problem with this approach is that it assumes that the Monte Carlo gamma-ray images correctly match the real gamma-ray images in every detail. To circumvent this uncertainty, a different approach was used to train the neural netwo ed in this work. No simulations were used. Instead, the training was carried out ixed signal and background sample (On-source data) and a pure background sa Off-source data). In place of the "classification score", the significance of the ence between on-source and off-source data was maximized. A 4::3::1 network sed, with 4 input parameters, 3 nodes in a single hidden layer, and one output no , giving a total of 15 weighted couplings and 4 thresholds. Adjustment of these 19 parameters effectively changes the boundary of the selected region in 4-dimensional parameter space.

To facilitate comparison with the "supercuts" gamma-ray domain, we the same four parameters as inputs to the neural network classifier: width, le distance and alp... The first two parameters ar... ...sure of the width and length of the roughly elliptical images, the 'distance' is a measure of the separation of the image centroid from the source position within ... field of view, and 'alpha' is the angle between the major axis of the ellipse and ... line joining the centroid to the source position. The "supercuts" hypercube wa... ...d as our starting point:

$$0.073° < \text{width} < \ldots$$
$$0.16° < \text{length} < \ldots$$
$$0.51° < \text{distance} < \ldots$$
$$0.0° < \text{alpha} < 15°.$$

All On and Off events within this region were fed ... the neural network, which then tried to find a better significance for the On-OFF difference by iteratively changing the neural network couplings to outline the best non-linear shape for the gamma-ray domain. Data obtained on the Crab Nebula in December 1993 using the Whipple Collaboration 10m Cherenkov Imaging telescope were used in the optimization. The database consisted of 19 On/Off pairs, with a total on-source time of 473 minutes. The standard "supercuts" effect was 16.6σ on this database; this increased to 17.8σ after optimization using the neural network. While this is not in itself a very marked improvement in sensitivity, it is nevertheless encouraging that some improvement is obtainable, particularly as the improved shape of the gamma-ray domain can be readily explained, as shall be discussed in the following section. Work is in progress on optimising a looser supercuts hypercube region, thus allowing for excursions of the boundary outside the supercuts domain as well as inside it.

3 VISUALIZATION OF THE NON-LINEAR GAMMA-RAY DOMAIN

A criticism sometimes leveled against the neural network approach is that is operates as a "black box" — the choice of values for the network couplings and thresholds cannot be explained on a physical basis. If, however, we can find a means of visualizing the resulting selected region in multiparameter space, and if we can ascribe physical reasoning to the shape of the region, then the network can be viewed simply as an efficient numerical approach to delineating a complicated non-linear boundary in multiparameter space.

The visualization approach used here is to look at density profiles on 2-dimensional projections of the multidimensional region. For an N-dimensional region, there are (N-1)+(N-2)+..+1 possible 2-D projections. Thus, for the 4-D parameter region in question here, we view 6 2-D projections. Fig. 1 illustrates the approach for the two simplest multiparameter shapes — the hypercu... ...nd the hypersphere. The region in question is first filled with a uniformly spac... ...sh of points, and the points are then projected and binned onto a 2-D plot. ...se shapes, all the 2-D projections look alike. Fig. 2 shows the 6 2-D projec... of the On-Off count differences within the supercuts hypersphere for the Crab ...la dataset. It can be seen that the distributions are highly non-uniform — i... were uniform, then no further optimization of the region would be possible. The plot of 'distance' against 'alpha' is of particular interest in the context of domain optimization. Most of the effect occurs at relatively large distances and relati... ...all alpha values. This is quite reasonable: gamma-ray images at large distanc... tend to be more elliptical, with well-defined major axes. Thus there will be less ...rtainly in the shower orientation

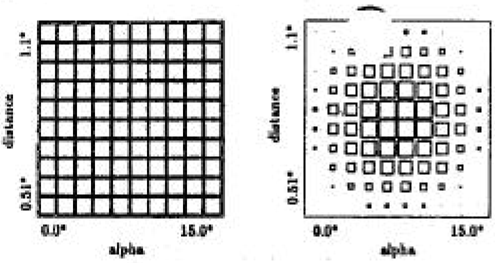

Figure 1: (a) 2-D projection of hypercube, (b) 2-D projection of hypersphere.

and the alpha distribution will be narrower for such events. This does not mean that rejecting all events with lower distance values or higher alpha values will lead to improved sensitivity; such simple linear cuts would give rise to another hypercube, and the supercuts hypercube is already close to optimal for this data.

To visualize the optimal non-linear region as selected by the neural network, we apply the network which was trained on the Crab Nebula data to a uniform mesh of points filling the supercuts hypercube. Projections onto 2-D plots will then show which edges of the hypercube have been eroded to improve the shape of the gamma-ray domain. The 6 2-D projections are shown in fig. 3. Not surprisingly, the new domain departs from linearity at high alpha values and large distances; from above, gamma-ray with large distance values should have tighter alpha values. Similarly, we observe a deficit in the region associated with large alpha values and large lengths; such events should have well defined major axes and thus tighter alpha values. Finally, we note a less pronounced deficit in the region associated with large alpha values and small width values; again, highly elliptical images will tend to have large lengths and narrow widths, so the same arguments apply.

We conclude that the neural network approach provides an efficient method of obtaining complex non-linear gamma-domain regions, revealing nontrivial structures in multiparameter space. Study of 2-D projections of the selected region show that it is physically reasonable, and thus helps dispell the impression that the neural network approach operates at a level which does not permit visualization or interpretation.

We would like to thank other members of the Whipple Collaboration for permission to use their Crab Nebula data in this study.

R[illegible]ERENCES

C[illegible]rian, A.A., 1991, Proc. 22nd ICRC, Dublin, 1, 540

F[illegible] D.J. *et al.* 1994, Proc. Workshop on "*Towards a Major Atmospheric Cherenkov De*[illegible] *II*[illegible] Ed. T. Kifune, Universal Academy Press, Inc., p149

H[illegible] A.M., and West, M. 1991, Proc. 22nd ICRC, Dublin, 1, 472

P[illegible] M. *et al.* 1991, Proc. 22nd ICRC, Dublin, 1, 464

Reynolds, P.T., 1991, Proc. 22nd ICRC, Dublin, 1, 496

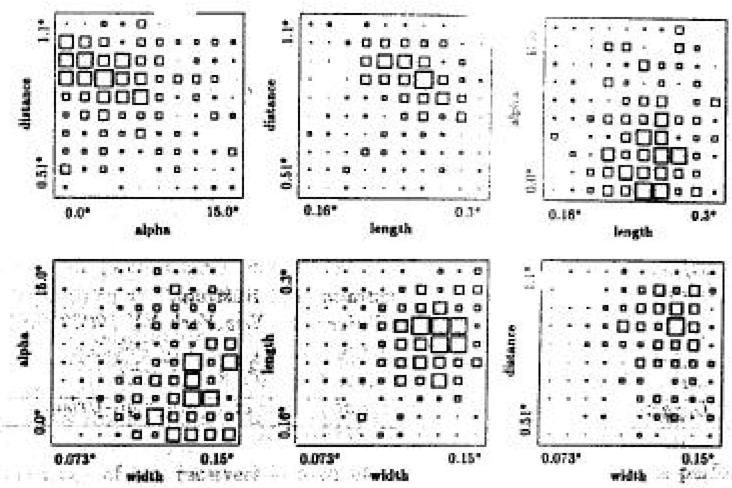

Figure 2: 2-D projections of Crab Nebula signal within Supercuts hypercube.

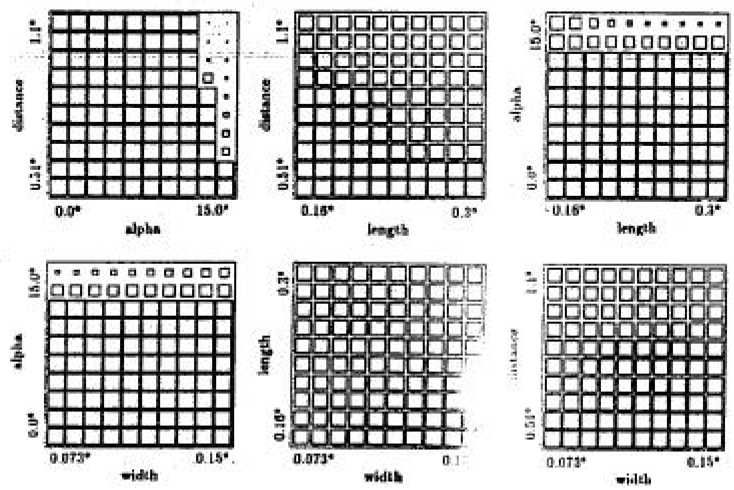

Figure 3: 2-D projections of region selected by Neural Network.