Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 131303
UNCLASSIFIED	

TITLE

Classification of passive sonar signals using a backpropagation neural network: simulation studies

System Number:

Patron Number:

Requester:

Notes: Paper #6 contained in Parent Sysnum #129006

DSIS Use only:

Deliver to: DK

#131303 #5

Classification of Passive Sonar Signals Using a Backpropagation Neural Network: Simulation Studies

Tim R.H. Cutmore
and
G. Robert Arrabito

Defence and Civil Institute of Environmental Medicine

Abstract

For the past several decades passive sonar signals have been processed by the human operator using visual and/or auditory displays to achieve detection and classification of target vessels. Recently, a small number of studies have begun to apply neural network techniques to this domain of signal processing. Time series data presents challenges to these pattern recognition networks and preprocessing methods are often critical. In this paper, a passive sonar hydrophone simulation is outlined. Methods and results of applying the backpropagation algorithm with a 3-layer feedforward network to simulated passive sonar data are presented. In particular, different types of fourier preprocessing and network parameters are examined for their effects on convergence rate (learning speed) and classification performance. Finally, the extension of this work to real passive sonar data and potential pitfalls are discussed.

1. Introduction

1.1 Background

The detection and classification of underwater acoustic sources are difficult and important problems, (Urick, 1983). The target sources may be masked by oceanic background noise and multiple targets may be present simultaneously. Classification, which assumes detection, has additional challenges of correct generalization and discrimination. For the past several decades passive sonar signals have been processed by the human-operator using visual and/or auditory displays to achieve detection and classification of target vessels.

Although the use of neural networks for processing sonar data is relatively recent, a large number of studies have been published for active sonar. These have concerned themselves with navigation of autonomous underwater vehicles (e.g., Shazeer et al., 1991; DeMuth et al., 1990; Schiller et al., 1989), classification of undersea objects (Gorman et al., 1988; Castellano et al., 1990; Porto, 1989; Venugopal et al., 1990) including mine classification (Bello, 1991). General views from a military perspective can be found in Webster (1991) and Hewish (1990). From a more esoteric perspective, a few studies have begun to examine mammalian echolocation with neural networks (e.g., Roitblat et al., 1989; in dolphins and Chen et al. 1988 in bats).

Work on passive sonar processing has been much less intensive. A recent search by the authors of several scientific data bases turned up only a few studies. Jacobson et al., (1988) conducted a preliminary study with a feedforward network trained by backpropagation to classify a set of natural acoustic sources derived from actual hydrophones. Various network configurations were examined and most were able to discriminate the target sources from each other and background noise. Liou (1988) using simulated data obtained partial success in classifying ship parameters, such as propeller shaft speed and number of propeller blades, with multiple sources present. Baran et al., (1991) used similar methods with natural data from hydrophone recordings of acoustic signals emitted by ships and was able to train a network to discriminate ship types. Bridle et al., (1990) applied a neural network in a sonar line tracking task.

1.2 Objectives

This report describes the initial steps towards complementary objectives: 1) to build a model of a natural acoustic source and 2) to develop in parallel a neural network architecture capable of discriminating features of this source. In building a model of an acoustic type complexity may be added by degrees to more closely approximate the natural source. This would permit a systematic study of the dimensions of the problem. For example, in the natural domain a single acoustic source may yield a variety of waveforms due to doppler effects and variation in the hydro-acoustic medium. Noise sources may mask the signal. In addition, several sources may be active simultaneously and this would require solving an acoustic occlusion problem.

In our selection of an acoustic source type it was of interest to compare machine verses human ability to classify sounds. To this end we have used a set of sources which were used by McFadden et al., (1992) in a study of how human subjects are able to classify acoustic sources (hereafter called "targets") and their associated "speeds". The quotations are used to indicate that these sources were simulated. The sound sources are described below.

1.3 Sound Simulation and Preprocessing

Figure 1 shows a system view. Sounds were synthesized using the CARL Cmusic software package (CARL, 1985). The McFadden et al., (1992) study used 10 target sources which were defined by a set of frequency components blended together. Three instances of each target were constructed by varying a "pulsing" component (simulated cavitation) which was interpreted to indicate the speed of the target. The sounds were presented in the auditory modality only (i.e., no visual displays were used).

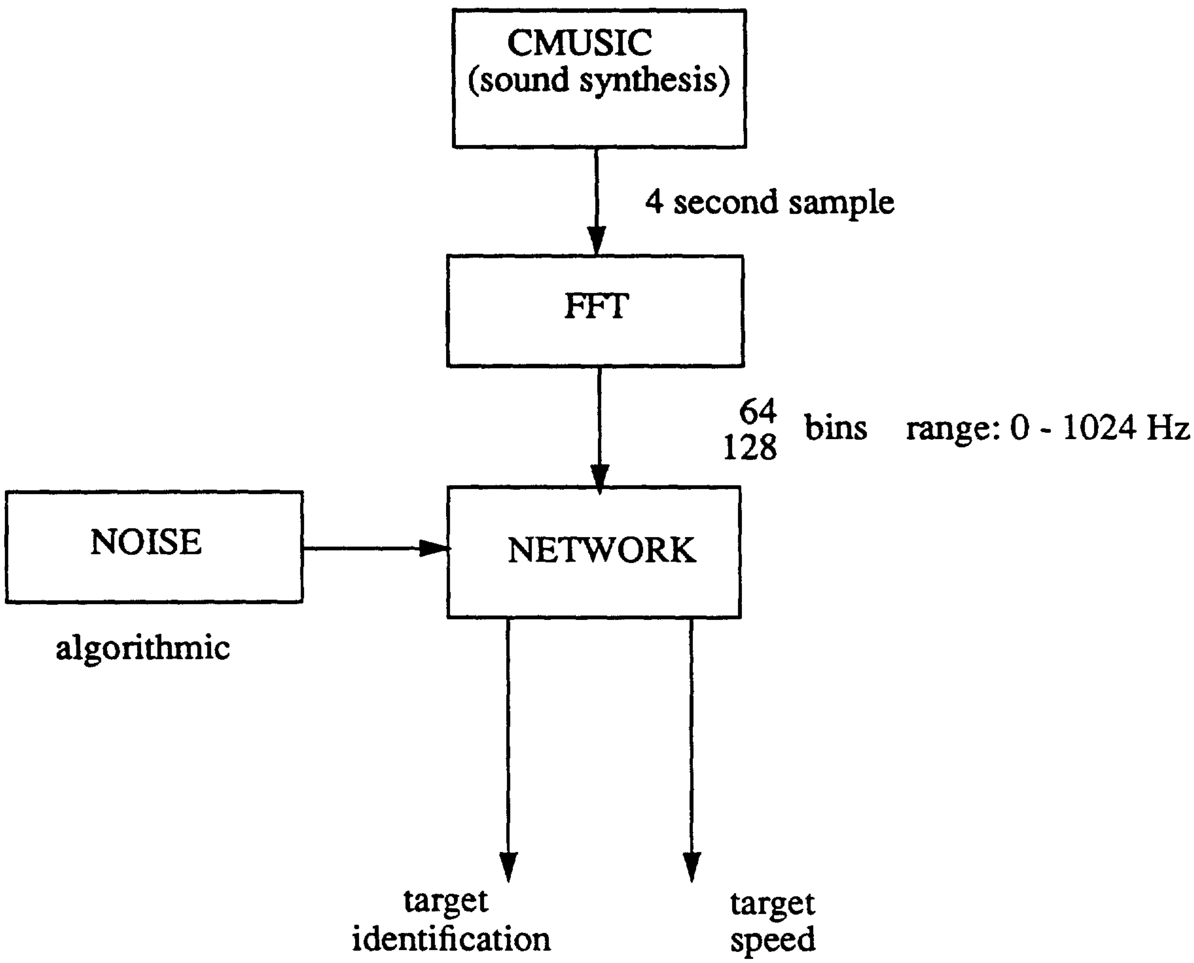


Figure 1

The sounds were passed through a Fourier Transform (FFT) to yield a power spectrum (as amplitude squared values) using a 4 sec sample at a sampling rate of 8192 Hz. Only frequencies below 1024 Hz were then retained, since none of the sounds produced power above this frequency. Two resolutions were examined, 8 Hz (128 power bins) and 16 Hz (64 power bins). Figure 2 shows examples of the 128 bin power spectra for each of the 10 targets.

Normalized FFTs of Targets

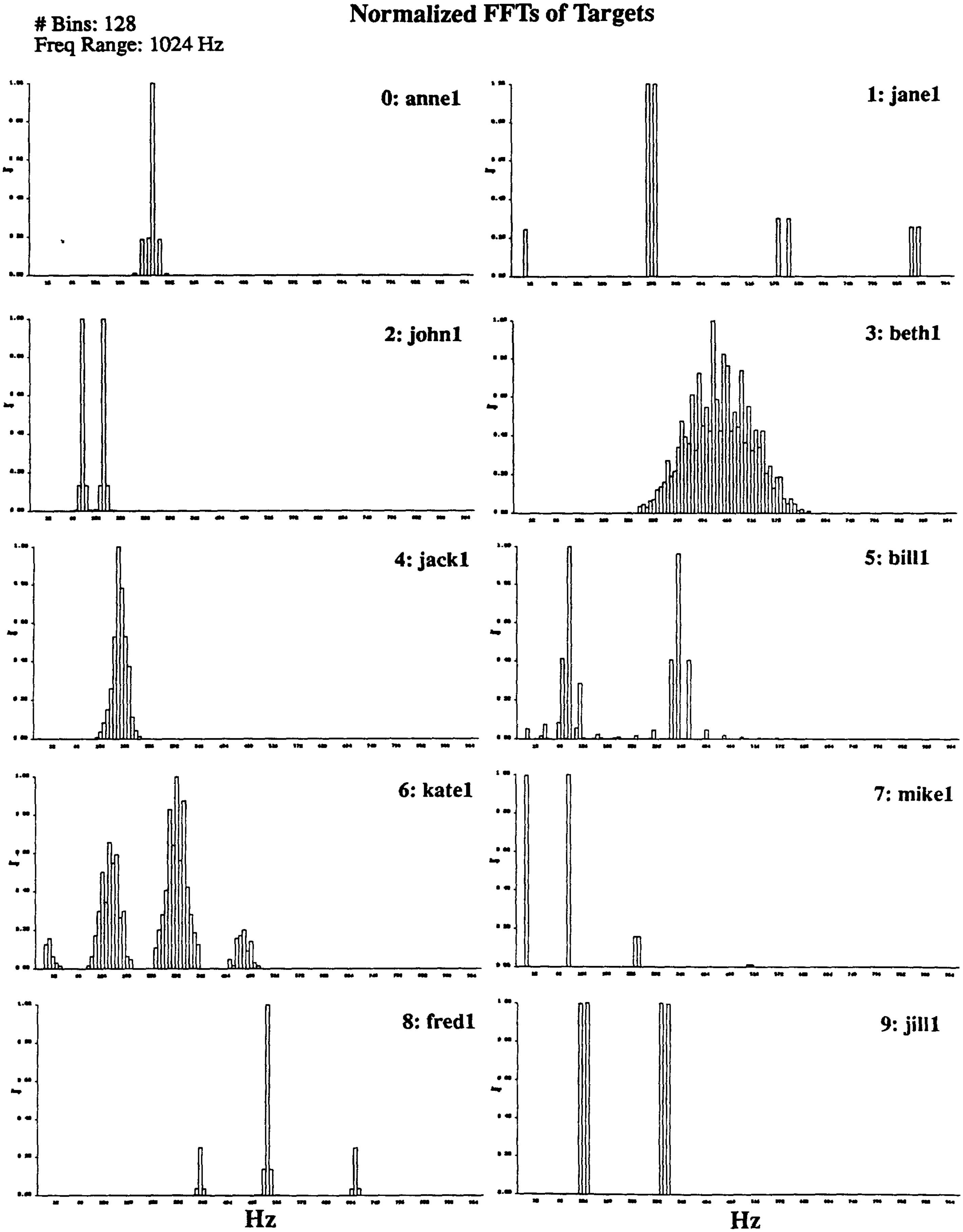


Figure 2

Figure 3 illustrates how the instances could vary for a given target, depending on the speed. Both simple additive low frequency components (pulses) and amplitude modulated cavitation was included. Twenty sounds (two from each target) were used to train the subjects (and the networks here) with the remaining 10 used to test generalization.

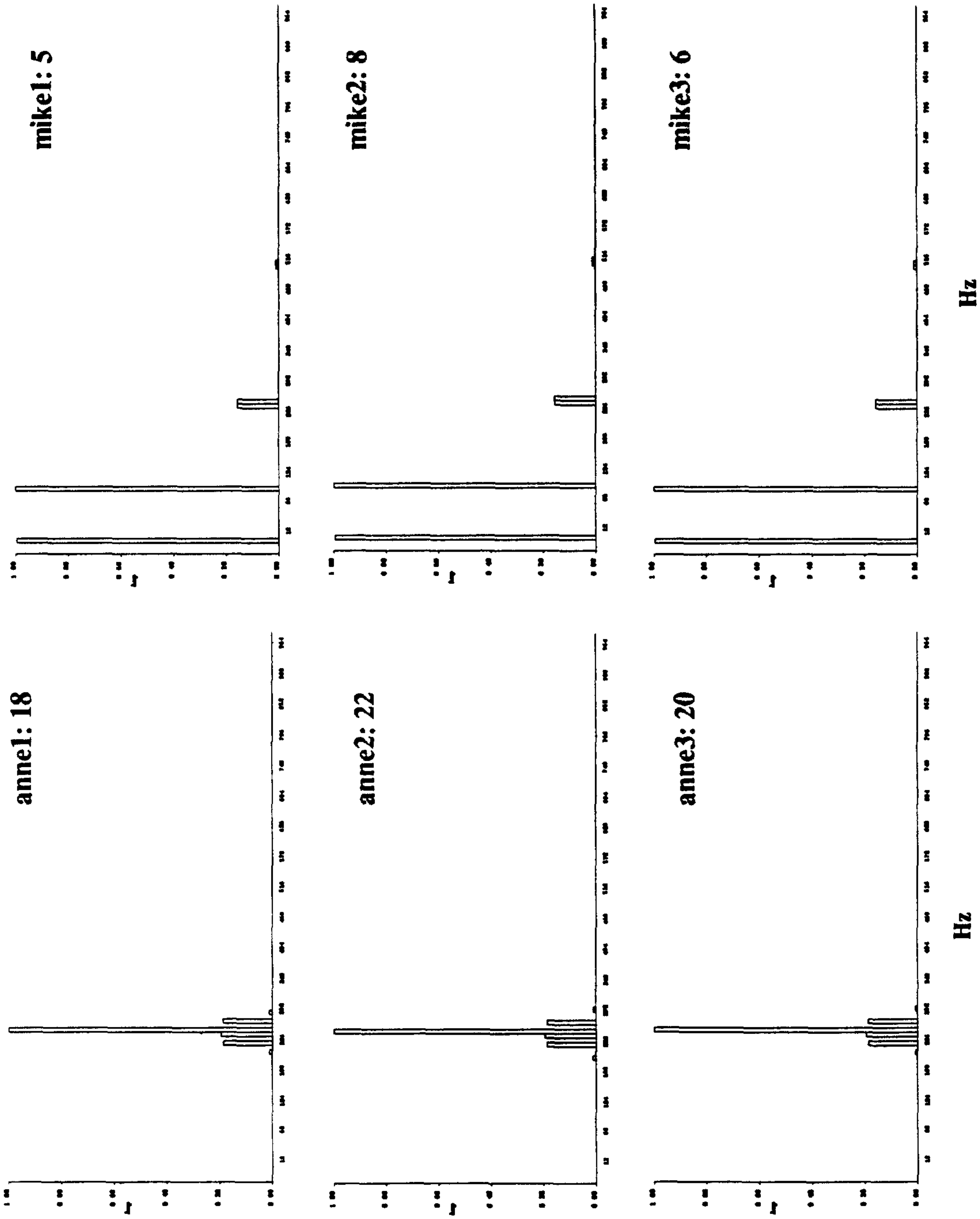


Figure 3

1.4 Network Selection

In selecting a simple feedforward network trained with backpropagation to perform the classification and speed estimation, its well studied structure, learning rule and associated parameters (momentum, temperature, learning rate) were decisive factors. If it should turn out that this extensively used network was inadequate for the task, then at least a baseline level of performance would be established for any future work with more elaborate networks. Other types of networks have been used to classify sonar signals. Genetic algorithms have been used to classify sonar images (Montana et al., 1989). Specht (1989) used an exponential in place of the sigmoid function to produce a "probabilistic network" which apparently can be trained much more rapidly than by back propagation. Good results were obtained with a sonar submarine detection task.

Figure 4 shows the network architecture that was used for both experiments. The inputs were the FFT normalized power bins. Various numbers of hidden nodes were used. Fifteen outputs were divided into two sets. One set was used to train the network as a classifier, with 10 alternatives (one for each target). The second set was used to train the network to produce a binary representation of the speed of the targets. A threshold was applied to both sets to determine whether an output node was "on" or "off". The activation function was a sigmoid with a range of -1 to +1. The generalized delta rule was used to train the network with a momentum term (Rumelhart et al., 1986).

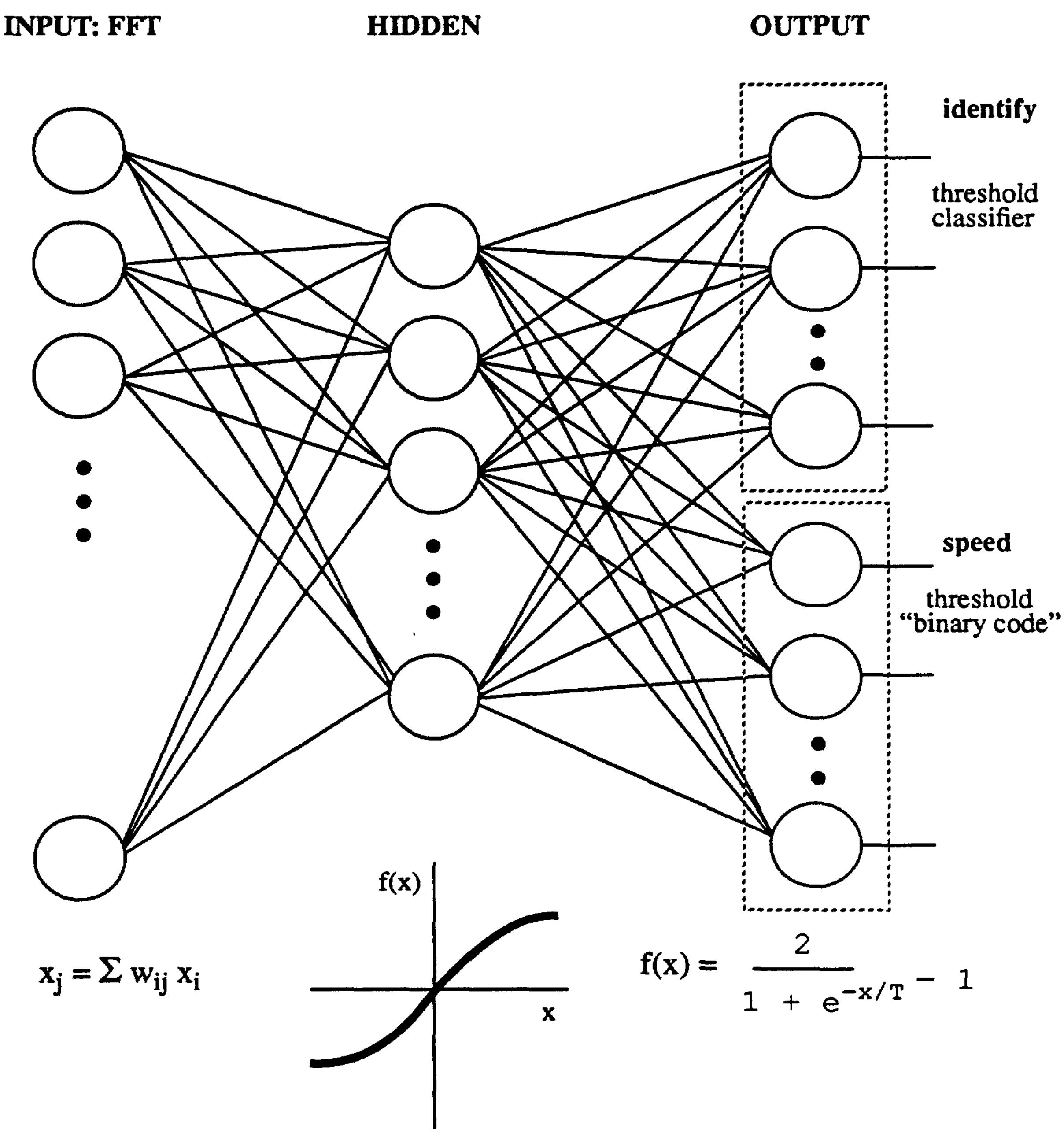
2. Experiment 1

The purpose of the first experiment was to examine a variety of parameters with "clean" (no noise) target signals in an attempt to find the "best" networks for solving the problem.

2.1 Network Designs

Figure 5 shows the parameters which were fixed and systematically varied. The fixed parameter levels were selected during a series of pilot trials. For example, it was found that an initial learning rate of 0.3 which was permitted to increase in proportion (0.4) to the change in RMS error at the output nodes, produced rapid, yet stable convergence. Bias inputs were generally not found to assist in convergence and so were left "off". Finally, a fixed number of cycles was selected at 2048, since it was found that in virtually all cases, convergence (if it was going to occur at all) was essentially complete at this time. The network was always trained by epoch. In other words, the full set of training patterns was presented before weights were updated.

FEEDFORWARD NEURAL NETWORK



$$x_j = \Sigma\, w_{ij}\, x_i$$

$$f(x) = \frac{2}{1 + e^{-x/T}} - 1$$

$$\Delta w = \lambda \delta x \; + \mu \delta w$$

$$\delta = (t - o)\, f' \qquad \text{(output layer)}$$

$$\delta_i = (\Sigma w_{ij} \delta_j)\, f' \qquad \text{(hidden layer)}$$

Figure 4

FIXED PARAMETERS:

1) Initial learning rate: $\lambda = 0.3$

2) Adaptive learning: a = 0.4

3) Bias Nodes: 0.0

4) Cycles: 2048

VARIABLE PARMATERS:

1) Temperature: T = 0.5, 1.0, 4.0

2) Momentum: $\mu = 0.0, 0.3, 0.9$

3) FFT Input Resolution: 8, 16 Hz (0 - 1024 Hz)

4) Hidden Nodes: 0, 4, 8, 16, 32

Figure 5

The range of the variable parameters was also selected following pilot work to assess their general efficacy. For example, extending the temperature beyond 4.0 produced similar results albeit with more learning cycles required. Temperatures below 0.5 produced very unstable behaviour. For momentum, values above 0.9 may produce unstable behaviour as well. Five variations of network were examined which differed according to the number of hidden units: none (perceptron), 4, 8, 16 and 32. These will be referred to as xH where x is the number of hidden nodes. As previously noted 8 Hz and 16 Hz FFT resolutions were used. The experiment thus included: three levels of temperature, three levels of momentum, two input resolutions and five hidden node alternatives. This produced 90 different network configurations. For each configuration, five replications were averaged, each having a different set of random initial weights. These weights were constrained between -0.5 and +0.5.

2.2 Results

The two measures of primary interest were percent of targets correctly identified and the error of speed estimation. The RMS error and learning rate measures receive brief comment.

The primary result of interest for this first experiment was the rapidity with which a given network could achieve high performance for identifying a target and estimating its speed, both for trained targets and for ones to which it had not yet been exposed. Therefore, the learning curves are shown to illustrate network behaviour.

Since there are a very large number of network configurations, illustrative sets are provided with some parameters ignored. These are divided into topics for emphasis: training and testing, temperature and momentum, and 16 verses 32 hidden nodes. Finally, 64 FFT bin inputs showed similar yet frequently slightly poorer performance for all measures, particularly for the speed estimation. For simplicity only the results with the 128 bin networks are shown.

2.2.1 RMS Error and Learning Rate: Figure 6 shows examples of the relationship between RMS error and learning rate. These (as well as all remaining graphs) show log linear plots of training cycles and a performance measure. The perceptron and 16 hidden node networks are shown with each of the T and μ levels. The RMS error for the test patterns are also shown. In general, networks were able to converge to an RMS error below 12% for at least some parameter combinations and the learning rate increased in proportion to this. However, as shown for the 16H network, T = 0.5 did not produce good performance. Finally, test patterns generally showed a minimum RMS error well above training patterns, at about 0.5 to 0.6.

2.2.2 Training and Testing: Figure 7 shows the results of training and testing for target identification for the 0H, 8H and 16H networks. Only $\mu = 0$ results are shown. The following findings were of interest here: 1) Each network has one or more parameter combinations that result in 100% correct target identification and the number of cycles required is exceptionally small - 4 to 64 cycles. This is the case for both training and test patterns as they nearly overlay each other. 2) The perceptron converges much more rapidly than networks with hidden nodes; 3) While the perceptron's training and test performance appears to be simply shifted to larger numbers of learning cycles as T increases, networks with hidden nodes have an optimal temperature.

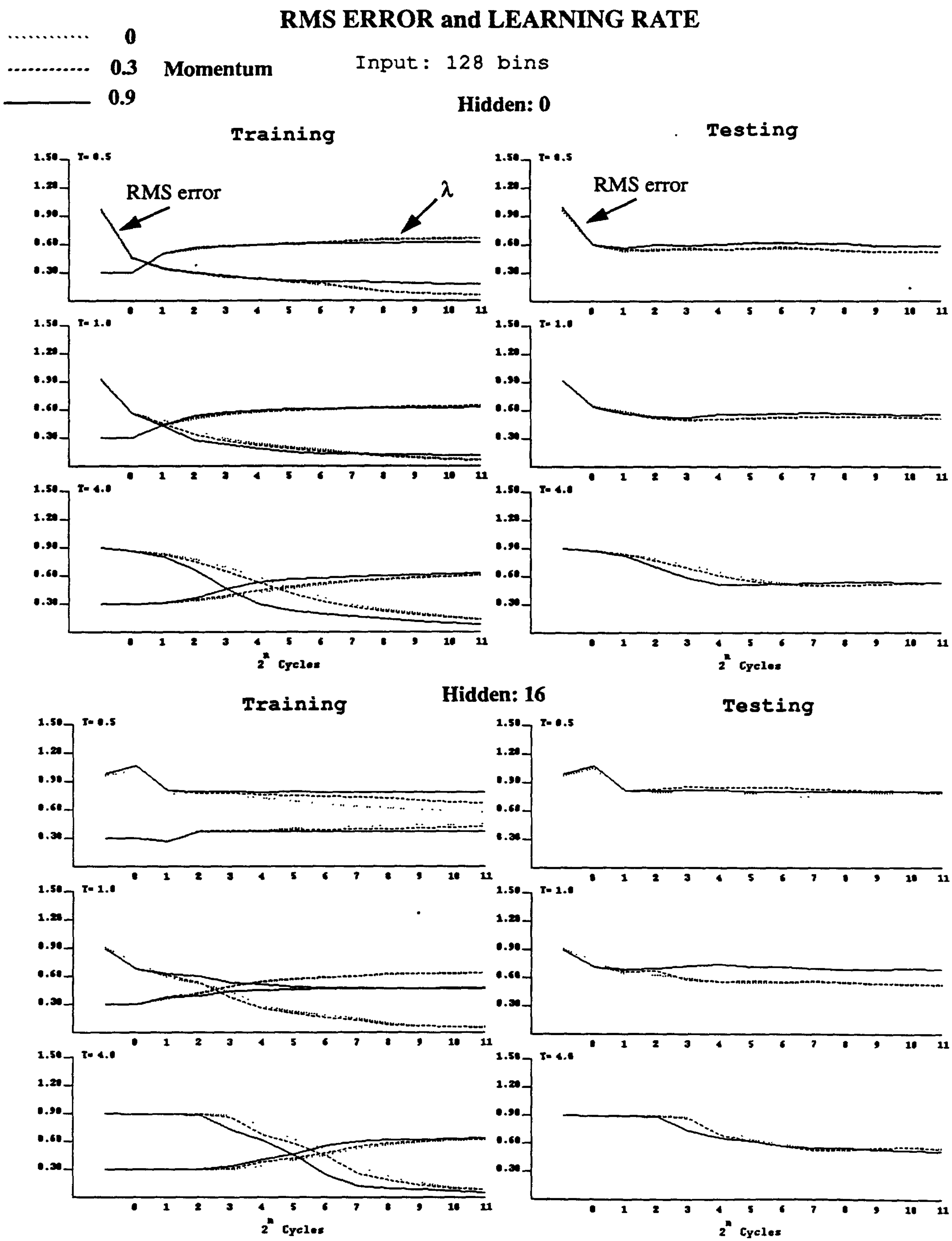

RMS ERROR and LEARNING RATE
0
0.3
Momentum
0.9
Input: 128 bins
Hidden: 0
Training
Testing
RMS error
λ
RMS error
T= 0.5
T= 1.0
T= 4.0
2^n Cycles
2^n Cycles
Hidden: 16
Training
Testing
T= 0.5
T= 1.0
T= 4.0
2^n Cycles
2^n Cycles

Figure 6

TRAINING and TESTING

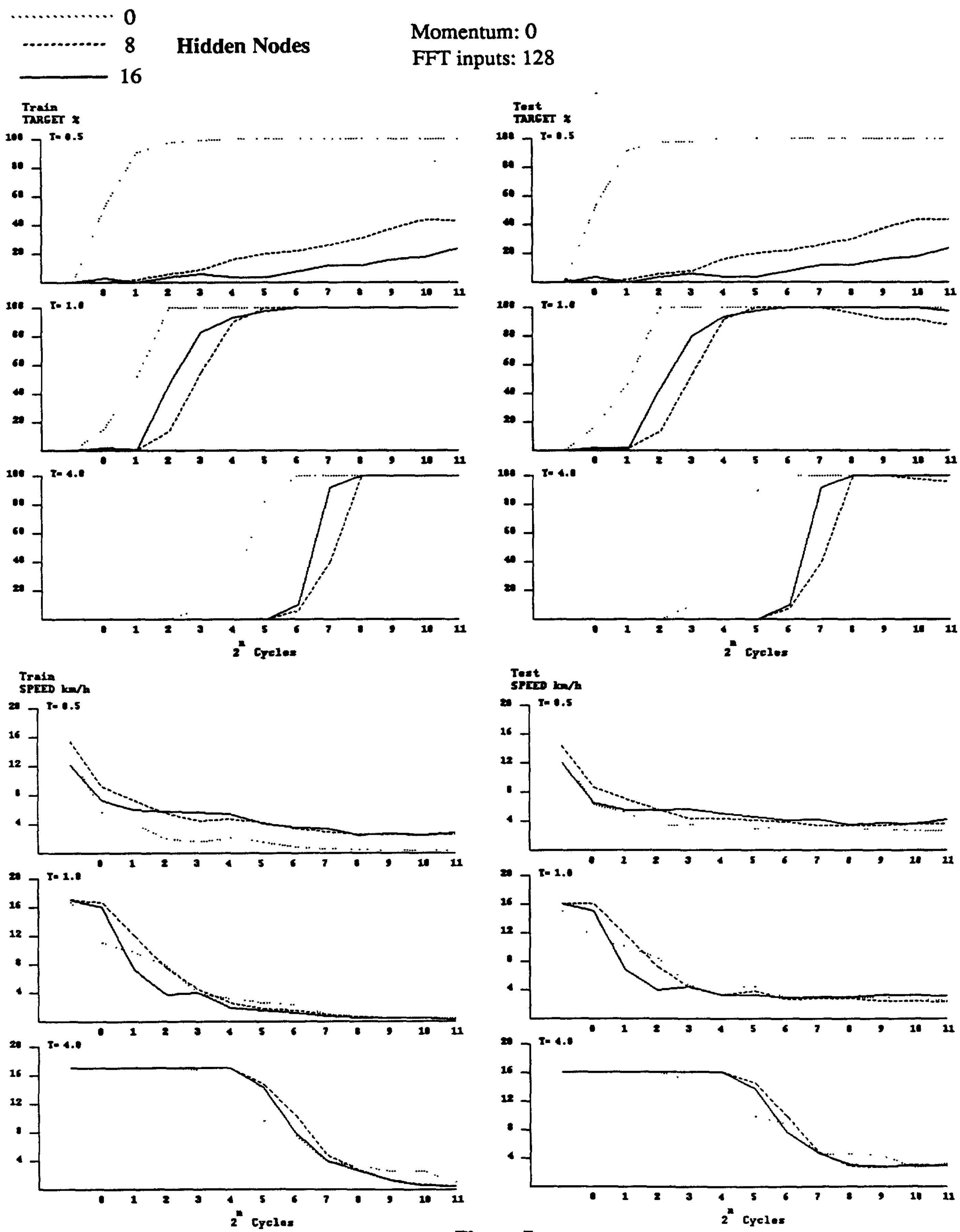


Figure 7

Figure 7 also shows the results for the speed estimation error. The results of interest are: 1) The error in the trained patterns approaches very near to zero, while the test patterns appear to yield at best a speed error of 4 units. 2) The perceptron is not so clearly the fastest learning network with this measure. Indeed at T = 1.0 the training and test performance of the 16H net was found to be equal to or better than all other networks. 3) A comparison of Figs 7 and 8 shows the speed estimation error requires more learning trials than target identification to achieve highly accurate performance for both training and test patterns. 4) An overtraining effect is evident for the 8H network at T = 1.0.

2.2.3 Temperature and Momentum: Figure 8 shows the effects of varying both temperature and momentum. From these data the "best" T and μ parameters were selected for each network class[1]. For the perceptron these were: T = 1.0; μ = 0.9; For the 4H network, T = 4.0, m = 0.9; For the 8H, 16H and 32H (not shown) networks, T = 1.0, μ = 0.3.

Figures 9 shows results which address the question of the utility of including 32 hidden nodes over 16 hidden nodes. All T and μ levels are shown. It appears that these networks perform very similarly for both training and test patterns. A slight advantage for the 32 node network appears for T = 4.0. Therefore, there may be little advantage to using more than 16 hidden nodes to solve the identification and speed estimation problems here.

2.3 Conclusions:

Trained Patterns

1. A network with two sets of outputs was able to learn both target classifications and speed estimates with very high accuracy. For some networks all 10 target id's were learned in as little as 4 iterations. In all networks, speed estimates took more iterations for good performance.

2. On both these measures, the network equals or outperforms human subjects who were trained to make auditory discriminations (McFadden, 1992). For target id Humans have shown a mean accuracy score of 65%, although one person did score 100%. Most of the network configurations learned target identification to 100% accuracy of classification.

---

1. The target test and both speed estimate results showed a similar pattern of results.

TEMPERATURE and MOMENTUM

FFT inputs: 128 (Training)

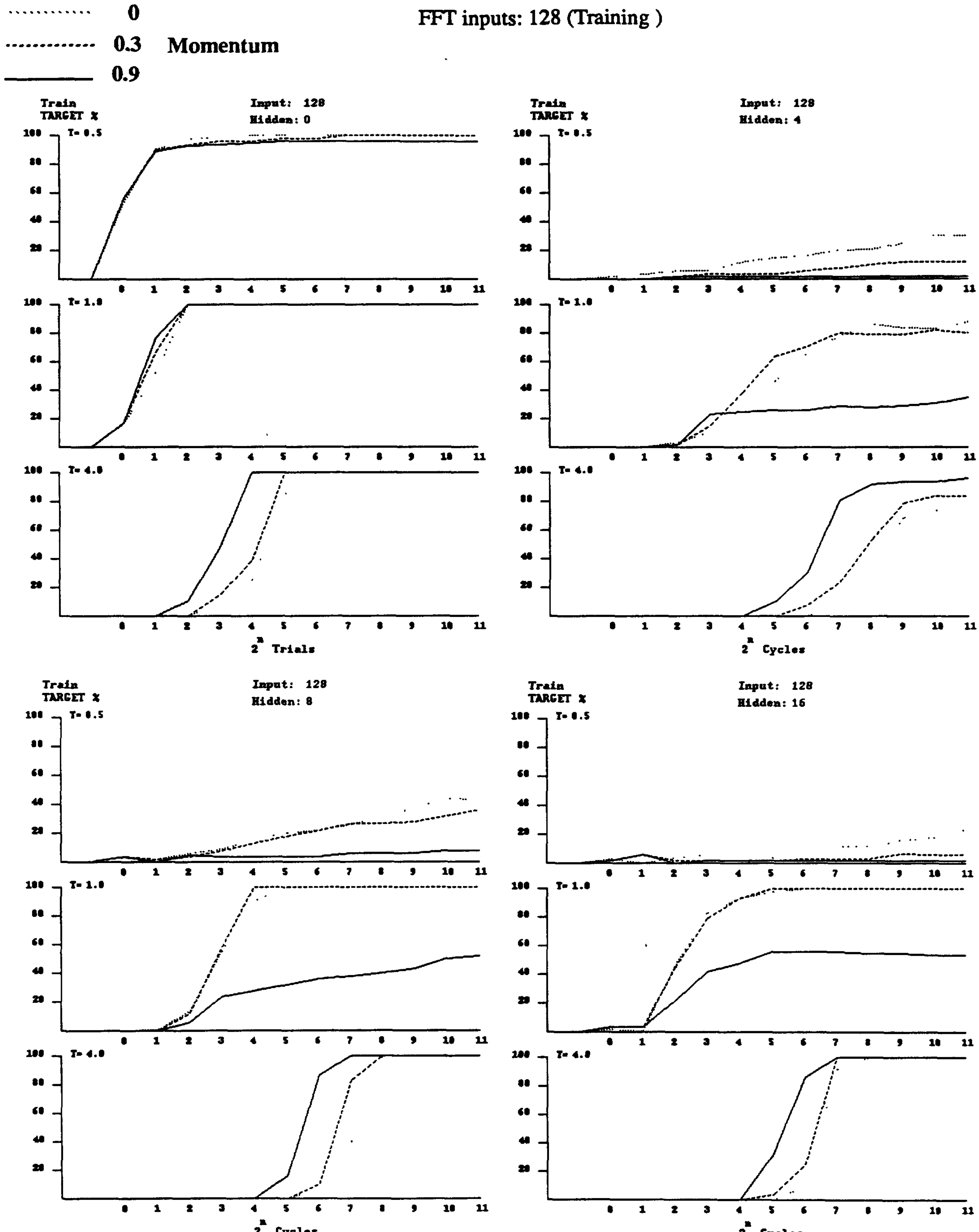


Figure 8

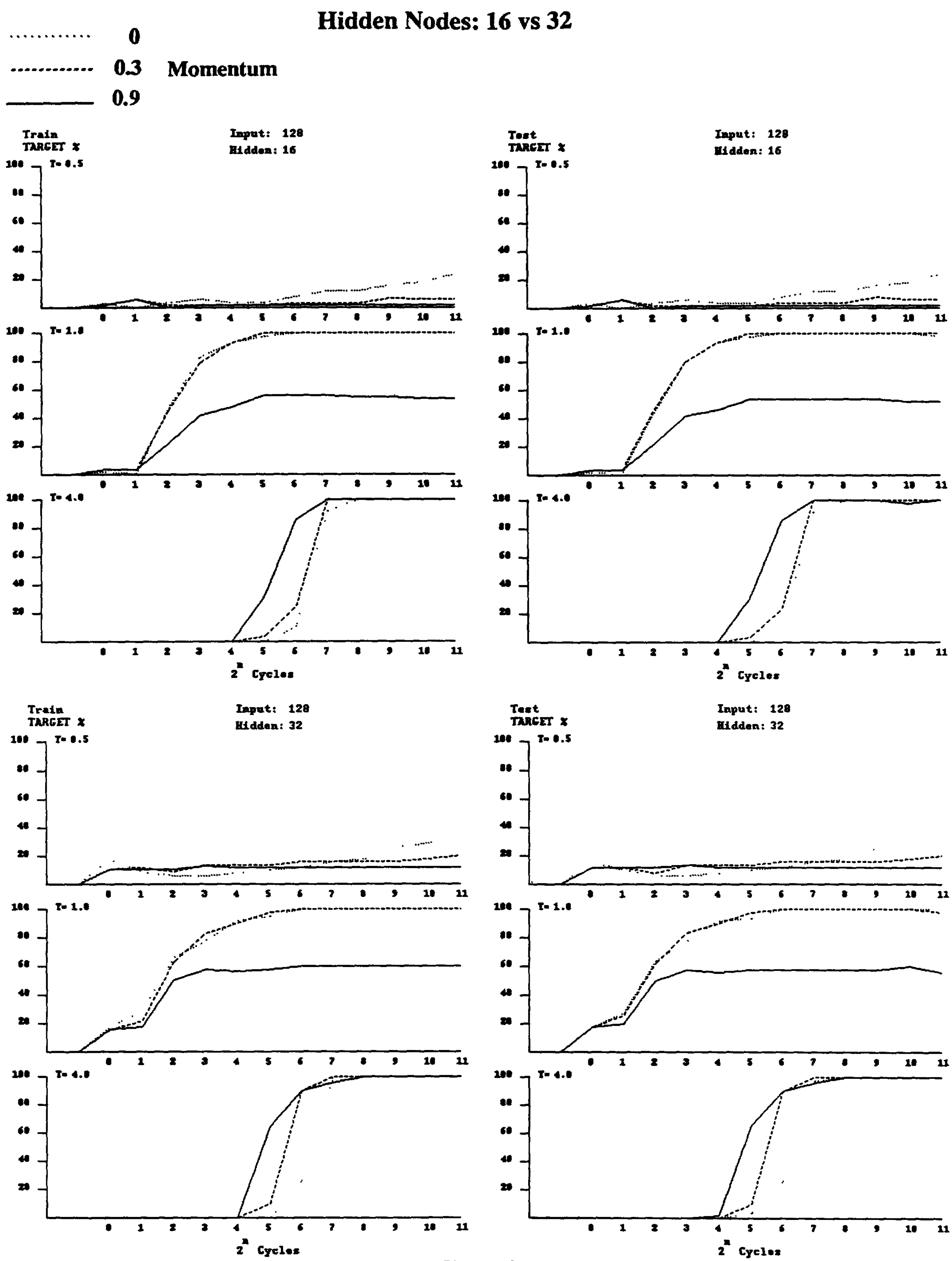

Hidden Nodes: 16 vs 32
0
0.3 Momentum
0.9
Train
TARGET %
Input: 128
Hidden: 16
Test
TARGET %
Input: 128
Hidden: 16
T= 0.5
T= 1.0
T= 4.0
2^n Cycles
Train
TARGET %
Input: 128
Hidden: 32
Test
TARGET %
Input: 128
Hidden: 32
2^n Cycles

Figure 9

age (since the normalized target signals varied in summed power the SNR's ranged from 2 db to -8 db). Two examples are shown in figure 10. A new noise pattern was generated each time a pattern was presented to the network. Two conditions were examined: 1) The noise was simply added to the inputs and the 10 classifier and 5 speed estimation nodes were used to train the network. 2) Everything in case 1 was repeated but in addition, an 11th classifier node was added to the output set and a novel noise pattern was presented to the network for each epoch of training (i.e. each cycle) and a unique noise pattern alone was presented for each test. Thus, the noise in both conditions was truly unique in every circumstance. The speed of a noise pattern was trained to zero.

These noise conditions addressed the issue of whether a feedforward network could handle both random variation in the input patterns as well as the question of whether the network could learn to recognize noise (or discriminate it) when it occurred in isolation.

3.2 Networks Used

From experiment 1, five networks (0H to 32H) were selected on the basis of best all around performance on classification and speed estimation. For 0H, $T = 0.9$; $\mu = 1.0$; 4H, $T = 4.0$; $\mu = 0.3$; for 8H, 16H and 32H, $T = 1.0$ and $\mu = 0.3$. In all cases the 128 node FFT input was used. The training and test pattern sets of experiment 1 were used for comparison to the clean target signal results. Again, five replications were averaged starting with different initial weights in each case. Finally, pilot tests with the noisy signals revealed oscillatory RMS error for the learning and adaptive parameters used in experiment 1. It was found that performance tended to be more stable for a low λ value. Therefore it was fixed at 0.1 with no adaptation.

3.3 Results

3.3.1 Target identification: Figure 11 shows the results for the five networks. The solid line is from experiment 1 for reference. The perceptron appears to have a notable problem for both noise conditions. For the noise added condition it does not achieve 100% target identification for either the training of test patterns. It does even more poorly for the noise discrimination condition. Since the graphs appear to reach asymptote after 512 trials, the perceptron may be incapable of learning to handle the noisy input patterns. The 4H network also had considerable difficulty. However, the figure suggests that, at least for training patterns, that continued

For the speed estimation error, humans learned to estimate to a mean score of 4 units. Most network configurations were able to learn this estimate to less than 1 unit.

Test Patterns (same target id's, different speeds)

3. Several networks classified targets (is) with perfect accuracy. Speed estimates were accurate to within 4 units for many networks.

4. On both measures, networks performed similarly to human subjects. For target id, mean performance for human subjects was at 65%, with the best scoring at 96%. Networks, in general, disseminated the test targets as well as the trained targets achieving 100% accuracy. For speed estimation, humans and networks scored similarly with an error of about 4 units.

Hidden Nodes

5. Observing the high performance of the perceptron the conclusion can be asserted that for a simple input space use a simple network (i.e., perhaps a perceptron will do). But as the input space become more complex, one or two hidden layers may become necessary.

6. Observing the difficulty with which the 4H network learned, it may be that such a network may have suffered from "overcompression" of information. However, it would be of interest to increase the learning cycles to determine whether better performance could be achieved.

Input Resolution

7. Fourier bin resolution of 8 Hz performed a little better than 16 Hz for resolving speed differences.

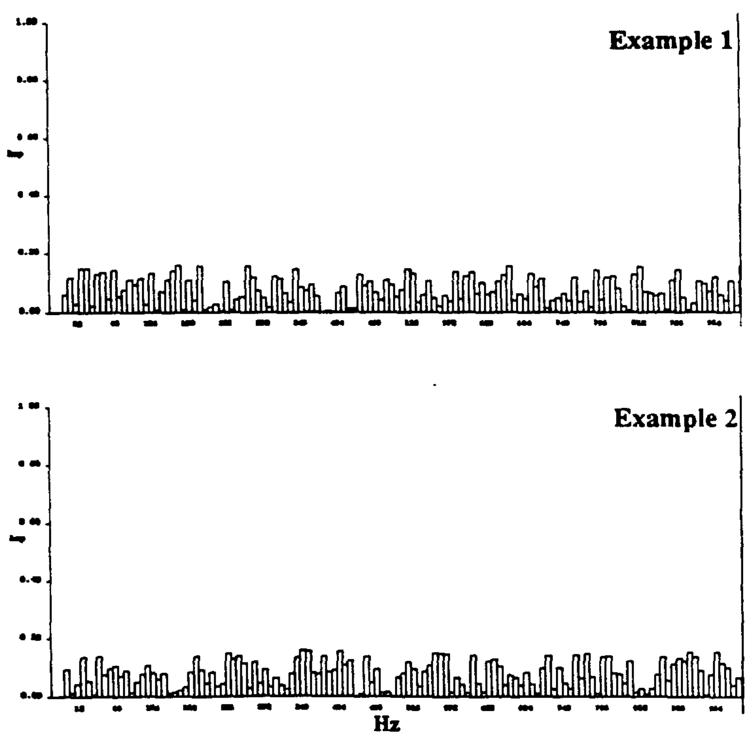
3. Experiment 2

Upon achieving the encouraging results in experiment 1 with clean signals it was of interest to examine the performance of the networks given a noisy input signal. This would improve the simulation model of natural target sounds in the ocean which has a broadband noise spectrum (Urick, 1983).

3.1 Noise Generation

Gaussian noise was digitized from a noise generator and added to the activation of the input nodes. The broad band (0 - 1024 Hz) signal to noise ratio (SNR) used was -5 db on avertraining beyond 2048 cycles may result in better performance, since the scores do not reach asymptote. The 8H, 16H and 32H networks performed progressively better. All three found solutions to achieve 100% recognition of trained patterns. Although the 8H and 16H networks achieved 100% test pattern recognition, this did not stabilize. The 8H network appeared to be particularly sensitive to over training as the decline in the noise added condition shows. Only the 32 node network learned both trained and test patterns to a stable 100% accuracy. In fact, this network even outperformed its processing of clean signals when noise was added to the patterns. This network achieved 100% accuracy after only 16 cycles with the noisy signals, but required 64 cycles with the clean signals.

SPECTRUM OF NOISE ADDED TO TARGET SIGNALS



Broad Band Signal to Noise Ratio: -5 db
Peak Single Bin Signal to Noise Ratio: 10 * log(1.0 / 0.2) = 7 db

Figure 10

NOISE ADDED TO TARGET SIGNALS
TARGET IDENTIFICATION

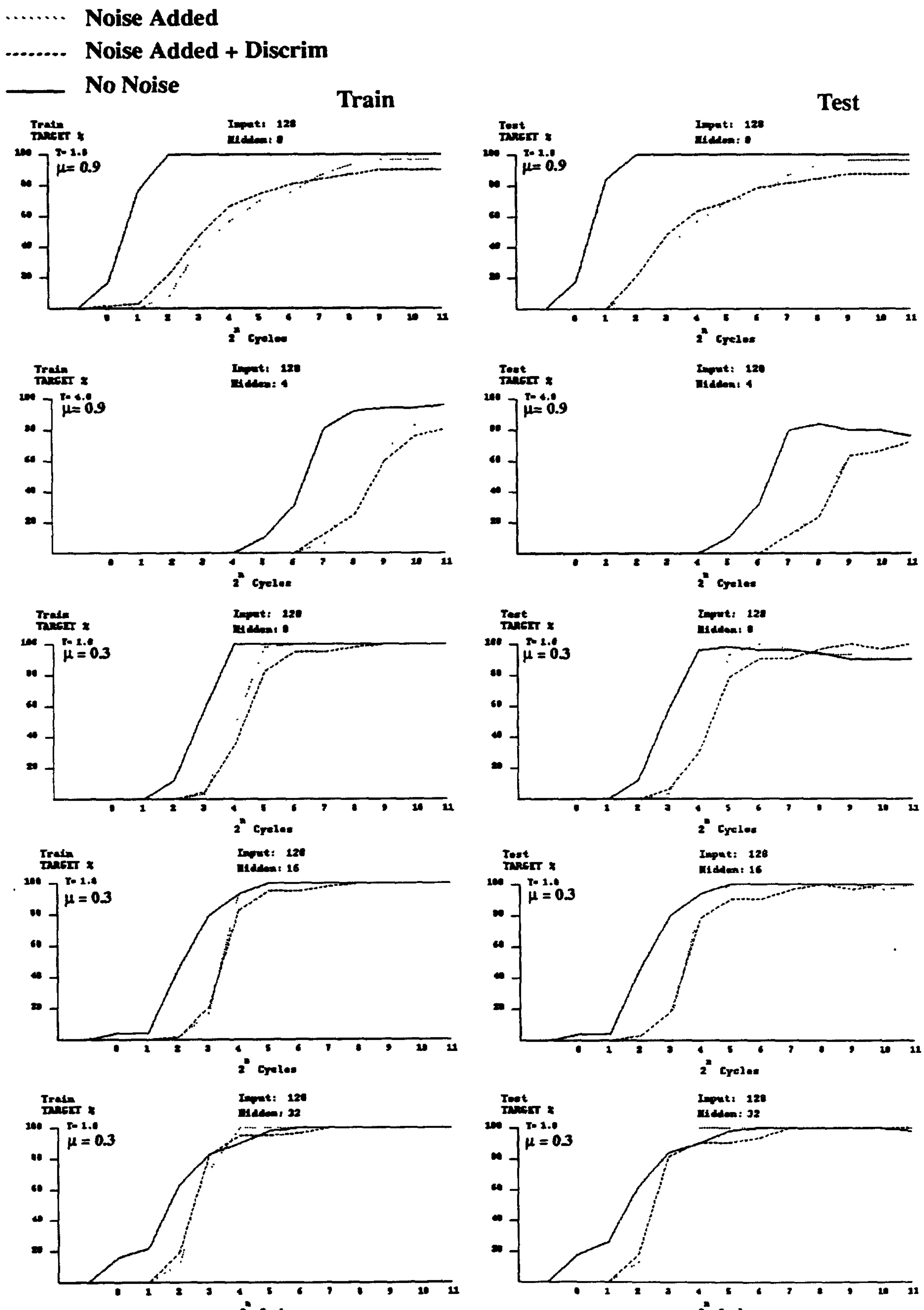


Figure 11

NOISE ADDED TO TARGET SIGNALS
SPEED ESTIMATION

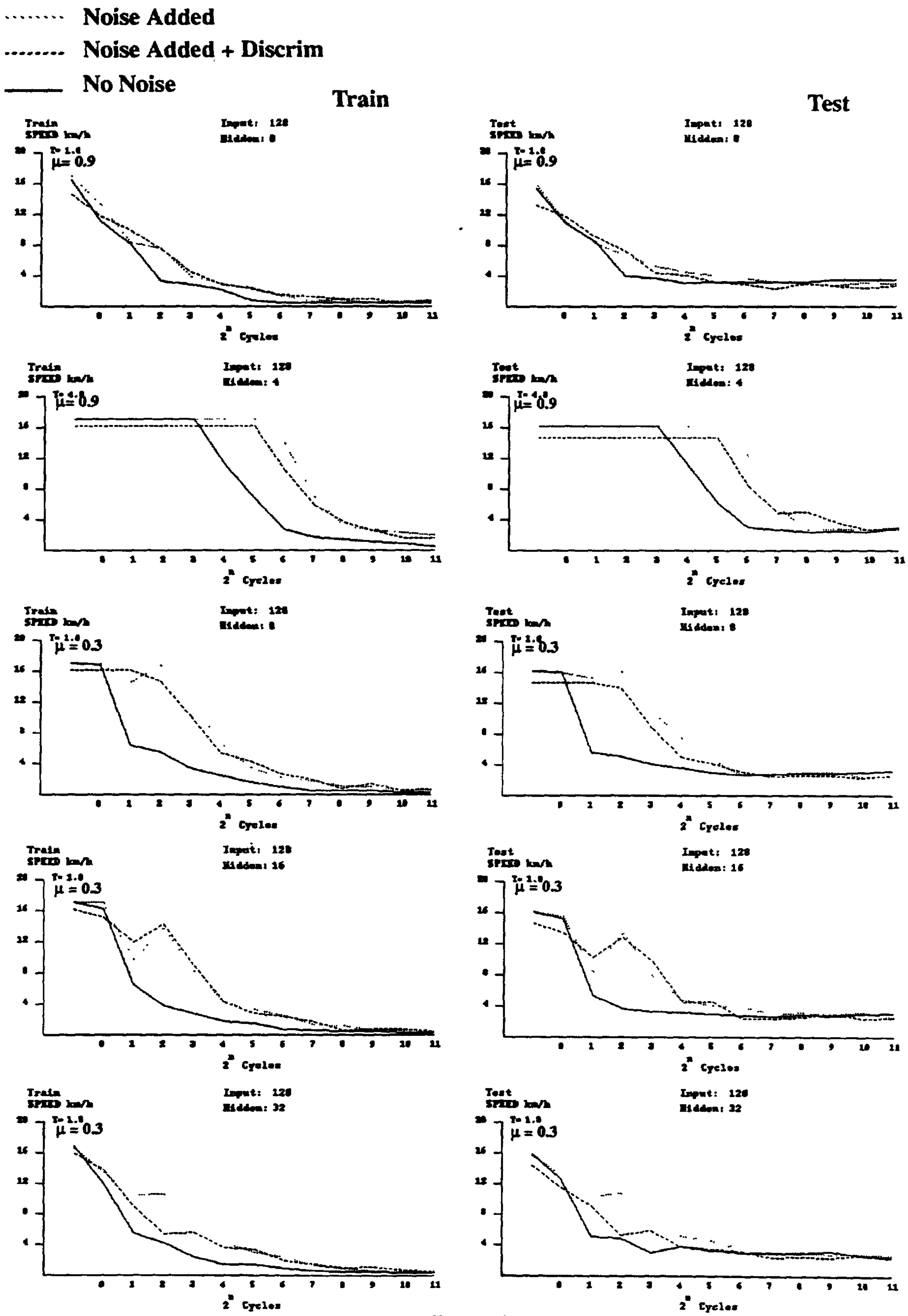


Figure 12

References

Baran, R.H. & Coughlin, J.P. (1991) A neural network for target classification using passive sonar. Proceedings of the International Society for Optical Engineering, 1471 (pp164-76).

Bello, M. (1991) Concepts for the use of three-dimensional depth maps in underwater mine classification. Report No. CSDL-R-2303, Charles Starck Draper Lab. Inc: Cambridge (51p)

Bridle, J. & Smith, D (1990) Line tracking using an artificial neural network and a dual process model Baysian mathematical framework. In: Colloquium on the Role of Image Processing in Defence and Military Electronics.

Computer Audio Research Laboratories (1985) CARL Startup Kit, University of California, San Diego.

Chen, L. & Brennan, C. (1988) Preliminary to a neural network model of sonar-based target discrimination in the echolocating bat. Tech Report, Providence, RI. Dept. Psychology(46p).

DeMuth, G. & Springsteen, S. (1990) Obstacle avoidance using neural networks. In: Proceedings of the Symposium on Autonomous Underwater Vehicle Technology, (pp213-215).

Gorman, R.P & Sejnowski, T.J. (1988) Analysis of hidden units in a layered network trained to classify sonar targets. Neural Networks, 1, (pp 75-89).

Hewish, M. (1990) Neural computing: Bringing brain power to defence. International Defence Review, 23: 2, (pp 17-19).

Jacobson, J.Z., Laplante, M. & Mogford, R.H. (1988) Artificial neural networks in naval warfare. Research Report 2-6803, Operations Research and Analysis Establishment.

Liou, C. (1988) Design of neural networks to classify sonar targets. In: International Neural Network Society (p 34).

McFadden, S.M., Belanger, J.M. & Yundt, K.A. (in preparation) Defence and Civil Institute of Environmental Medicine.

Montana, D.J. & Davis, L. (1989) Training feedforward neural networks using genetic algorithms. Proceedings of the 11th International Joint Conference on Artificial Intelligence, (762-7).

Porto, V.W. (1989) Pattern recognition of undersea objects using a neural network. In: 23rd Annual Asilomar Conference on Signals, Systems and Computers, (pp 376-380).

Roitblat, H.L., Moore, P.W.B., Nachtigall, P.E., Penner, R.H. & Au, W.W.L. (1989) Dolphin echolocation: Identification of returning echoes using a counterpropagation network. In: International Joint Conference on Neural Networks, (pp 295-300).

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986) Learning internal representations by error propagation. In: Parallel Distributed Processing Eds. D.E Rumelhart & J. L. McClelland. Cambridge: MIT Press.

Schiller, I. & Tench, K.A. (1989) A neural-network-based autonomous underwater vehicle guidance system. Proceedings of the 6th International Symposium on Unmanned Untethered Submersible Technology, (pp 312-19).

Shazeer, D.J. & Bello, M. (1991) Detection and classification of undersea objects using multilayer perceptrons. In: Applications of Artificial Neural Networks II (pp 622-636).

Specht, D. F. (1989) Probabilistic neural networks (a one-pass learning method) and potential applications. In: Wescon /89 Conference Record, Electronic Conventions Inc. (pp 780-5).

Urick, R.J. (1983) Principles of Underwater Sound, New York: McGraw-Hill.

Venugopal, K.P., Pandya, A.S. & Sudhakar, R. (1990) Sonar target recognition using neural networks, 3rd International Workshop. Neural Networks and Their Applications, (pp 563-565).

Webster, W.P. (1991) Artificial neural networks and their application to weapons. Naval Engineers Journal, 103, (pp 46-59).