A Counterpropagation network model to recognize and classify chart patterns in automated manufacturing

D. T. Pham[1] and Z. Yilbas[2]

[1] School of Electrical, Electronic and Systems Engineering, University of Wales, College of Cardiff, P.O.Box 904, Cardiff CF1 3YH, U.K.

[2] Erciyes Universitesi, Muhendislik Fakultesi, Elektronik Muhendisligi Bolumu Kayseri, TURKIYE.

Abstract

The present study investigates applicability of the counterpropagation networks to recognise and classify the six chart patterns which may be produced from the industrial or/and commercial analysis. Counterpropagation networks produced initially fails to accomplish this goal due to overlapping of some of the chart patterns in data and test files. In order to avoid overlapping and complexity of the chart patterns, the data file is divided into sub-data files in which six chart patterns are included. Ten counterpropagation networks were trained using this sub-data files. Five of the networks having high performances are selected. A combined network is constructed using this five networks set and one other sixth network. The combined system then tested and it is found that the combined network system introduced in the present work gives improved performances for the chart patterns.

1. INTRODUCTION

Counterpropogation network was invented by Robert Hecht-Nielsen of Hecht-Nielsen Neuro-Computer Corporation[1]. The original basis for the network is not clearly known, but it appears to have developed as a means for synthesizing complex functions, much as backpropagation. It works as a "look-up" table, in parallel finding the closest example and reading out its equivalent mapping.

Counterpropogation network selects from a set of exemplars by allowing them to compete amongst each other. Normalized inputs and competition between exemplars the nearest neighbour. This provides a method of constructing an adaptive pattern classifier, function approximation and data compression [2].

The counterpropagation network developed formerly consist of bi-directional mapping network. X and a Y inputs are applied, which in turn interact together to select a particular processing element which is nearest to the composite vector (X,Y). The winner activates X' and Y'. In this way it acts as an be-lateral auto-associative network. The network gets its name from the counter-posing flow of information through it.

Counter propagation networks are usually trained to perform pattern mapping, the mapping of one particular pattern to another for entire set of patterns. The trained network classifies that pattern into a particular group by using a stored reference vector. This provides the target pattern associated with the reference vector as a output. In this case hidden layer performs a competitive classification grouping the patterns. It was shown that, counterpropagation works best when the patterns are tightly clustered distinct groups [3]. It was also shown that counterpropagation may provide an excellent example of a network that combines different layers from other samples to construct a new type of network [4].

Two types of layers different than each other are used in counterpropagation. This includes: i) The hidden layer is a Kohenen type, in this case competitive units accomplish unsupervised learning. ii) The top layer is the Grossberg layer. In this case this layer is fully interconnected to the hidden layer and is not competitive. The Grossberg layer is trained by a Widrow-Hoff or Grossberg rule.

Counterpropagation is considered a faster alternative to backpropagation. The improvement in the training time with counterpropagation is in general substantial. Counterpropagation can learn many pattern mapping problems well but, recent studies have shown that it often generalizes less well on new patterns [5]. It has been shown that counterpropagation requires that input pattern classes be organised into clusters that are seperated (non-overlapping) [6]. Therefore in the light of the previous work [5,6] the present study examines the recognition and classification of the six chart patterns ,which are introduced in data files, using counterpropagation algorithm. To achieve this goal, a variety of counterpropagation networks, which has different number of processing elements in their hidden layer, i.e. (60-35-6), (60-30-6), (60-9-6) and (60-7-6) are trained using the data file developed before [7]. The counterpropagation networks trained are tested by using test file and their performances are produced. However, the performance of each networks fails to recognise all the chart patterns. This may be due to the fact that the pattern classes are not clearly seperable in the data file. In order to overcome this tackle, the data file is divided into sub-data files in which chart patterns are presented avoiding the overlapping problem. This provides six chart patterns taking place in one sub-data file. Ten different counterpropagation networks are trained using ten different arbitrary selected sub-data files. Performances of these networks are examined for individual chart patterns.Two sets of five networks having better performance than the others are selected. Each set of five networks are combined to the sixth network which in turn produces combined network system, i.e. two combine network systems each having six counterpropagation networks are produced. These two combined network systems are then tested by using the all sub-data files. Finally, the performances of these combined networks are examined and hence on of the combined network system is selected due to its better performance.

2. NETWORK MODELLING

The data file developed previously [7] are divided into sub-data files such that each sub data file contains the data corresponding to six chart patterns. i.e. each sub-data file has 6x60 data. Figure 1 shows control chart patterns. Each network is trained and tested using with one of these sub-data files. This provides very high level learning (100%) for each network. After completing the training, some of the weak connections between Kohonen and the output layer of the network are deleted. After this process the simplification of the network is obtained. This procedure is repeated for the other networks.

In order to construct and analyse the combined network system, each network is studied individually. In the combined system, five networks are employed and each of them has 75 processing element (60-9-6). Their first layer has 60 processing elements and it acts as a buffer. The input vectors are "normalized" and have same length. There are 9 processing elements in the competitive (Kohonen) layer different number of processing elements are experimented and 9 processing element in the Kohonen layer are found to be suitable for this application. In this layer processing elements are computed and one of the highest output wins. For a given inputs, one and only one Kohonen neuron output is a logical one and all others are zero. 6 processing elements take place in the output layer. Output layer provides a way of decoding that single input to a meaningfull output class. Widrow-Hoff learning rule is used in this layer. The output layer fully connected to Kohonen layer, the Kohonen layer fully connected to the prior layer. In this case, it has same effects as the Grossberg Outstar learning rule, but easier to implement. Initially, ten networks were trained seperately, by using 360 (6x60) different data. Each network learned 100%. It should be noted that each network was trained with using one of the different arbitrary selected sub-data files. The performance of the each network can be defined as:

$$\text{Performance} = \frac{\text{Number of Patterns Correctly Classified}}{\text{Total Number of Tests for the Pattern}}$$

The combined network system consist of 6 networks; five working simultaneously to process the input data and producing output data to feed the sixth network. Consequently, the output of the sixth network is the output of the combined network system. The block diagram of the arrangements of the networks in the combined network system is shown in figure 2. To obtained a high accuracy for the combined network system two sets of networks are selected

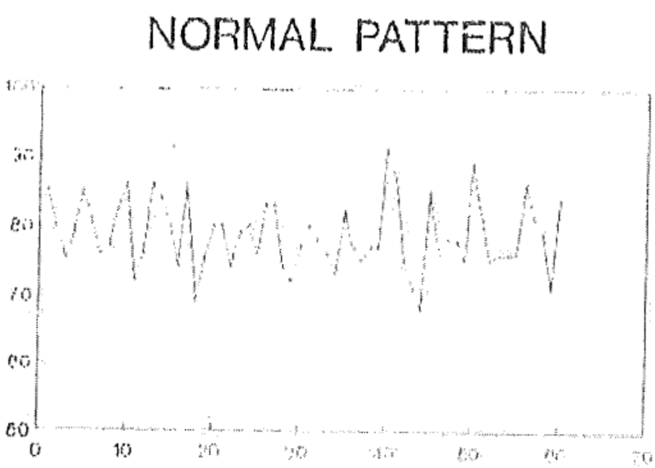

Figure 1.a. Pattern Number 1.

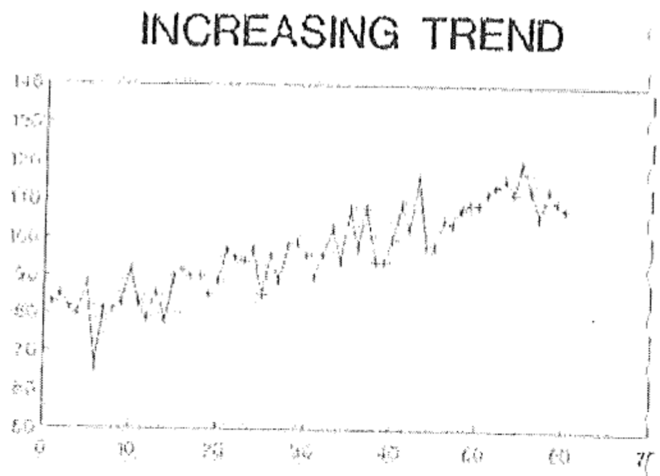

Figure 1.b. Pattern Number 2.

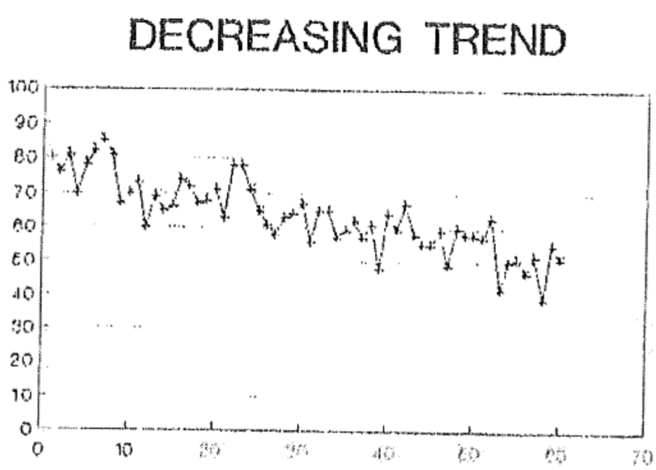

Figure 1.c. Pattern Number 3.

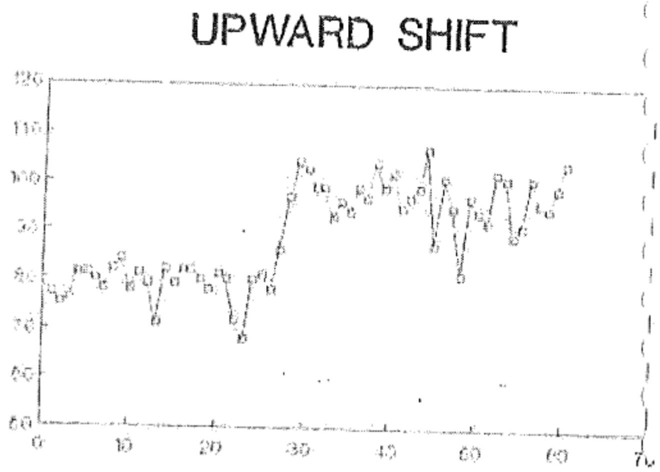

Figure 1.d. Pattern Number 4.

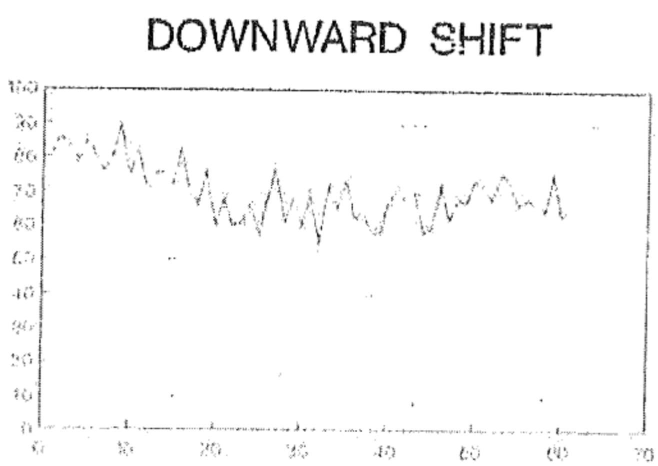

Figure 1.e. Pattern Number 5.

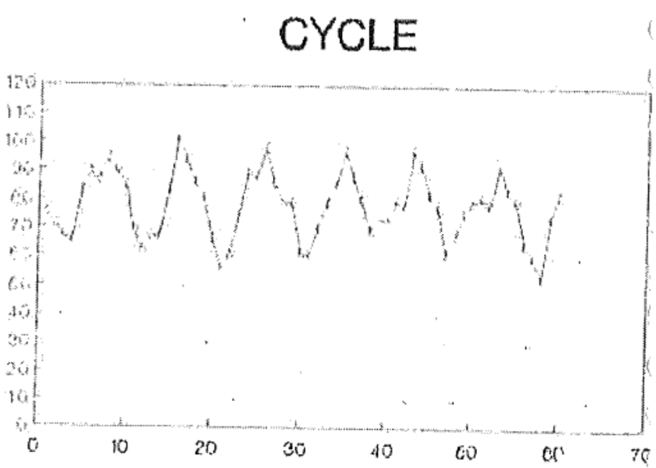

Figure 1.f. Pattern Number 6.

initially and their performances are compared. After the comparison, the set of networks having higher performance than the others are selected and this is employed in the combined network system. Finally, a data file is developed in a binary form to train the sixth network in the combined network system (Figure 2).

The sixth network in the combined network system has 42 processing elements (30-6-6). 30 of them are used in the input layer, 6 processing elements are used in the competitive (Kohonen) layer and last layer is the output layer with 6 processing elements. The data file for the sixth network contains 180 data -corresponding to 6 control chart patterns- obtained from the output of the set of five networks.

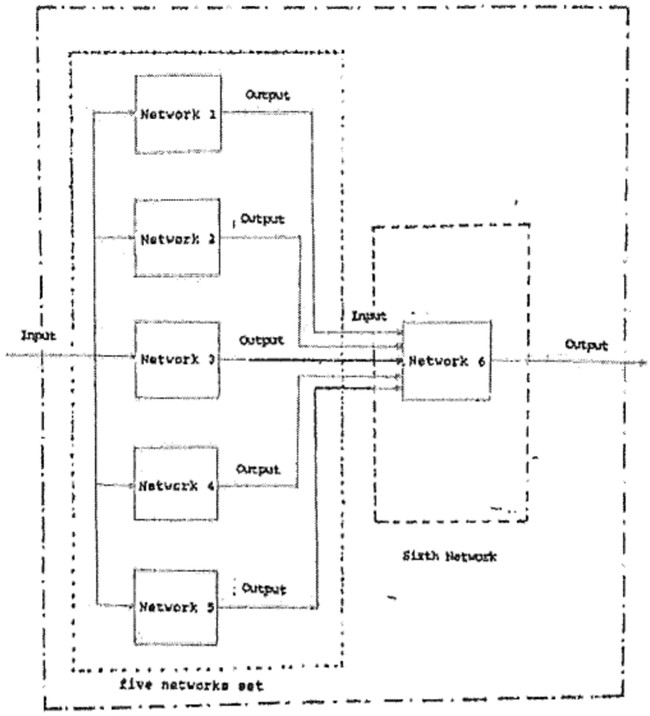

Figure 2. Block Diagram of The Combined Network System.

3. RESULT AND DISCUSSION

Each of the network in the combined network system is trained 3300-3400 times. After this training each of them learns 100%. So if the training times between counter-propagation and backpropagation is compared, counter-propagation nets learn faster than backpropagation nets. They are typically 10-100 times faster to train than conventional backpropagation; with the results that often comparable. The use of counterpropagation networks also provided rapid prototyping of the system. On the other hand, the hybrid schemes are not optimal in the sense that backpropagation is, since the hidden layer responses are not optimized with respect to the output performance. Figure 3 shows 3-D plot of performance with number of control chart patterns and number of network while figure 4 shows variation of performance with number of control chart pattern for ten networks. Performance of the networks varies with varying control chart patterns. This is due to the fact that each control chart pattern represented in different test files varies and networks are tested with the file different than their training files, which in turn gives variation in the performance. Two sets of five networks having high performance corresponding to all control chart patterns were selected from Figure 4.

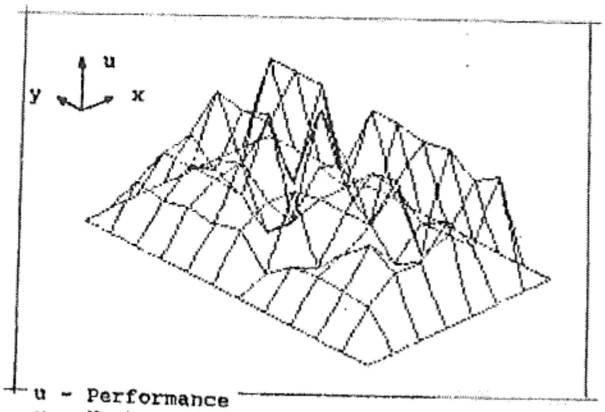

Figure 3. 3-D Plot of Ten Networks Performances With Chart Patterns.

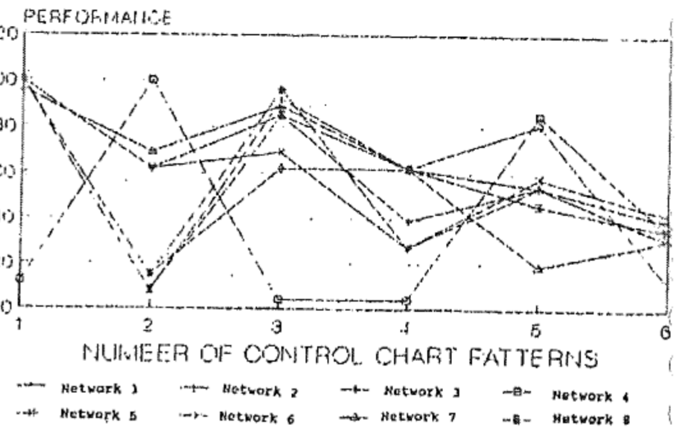

Figure 4. Variation of Performances of Networks With Chart Patterns.

Figure 5 shows the performance of the two combined network systems, providing that each combined network system employes a set of high performance five networks system selected before. The performances obtained from two combined network systems are almost the same. However, one of the combined network system corresponding to set-1 has slightly better performance than the other. Consequently, this combined network system is selected. It may be seen from Figure 5 that performance of the combined network system selected is high for the pattern numbers 1, 2, 3, 5 and it is especially low for the pattern number 6. In order to investigated the performance of each combined network system, the arithmetic mean value of the performances corresponding to each of five networks used in the combined system are plotted together with performance of the combined network system in figures 5.a and 5.b for the control chart patterns. It is obvious that the performance of the combined network system is higher than the aritmetic mean value of the performances corresponding to set of five networks. This clearly indicates that combined network system gives higher performance than the other trained networks.

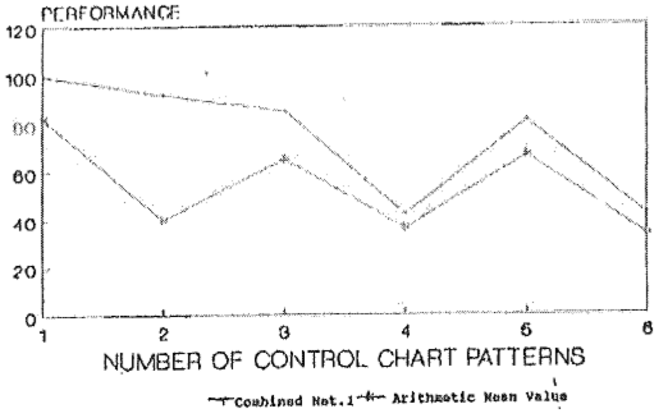

Figure 5.a.

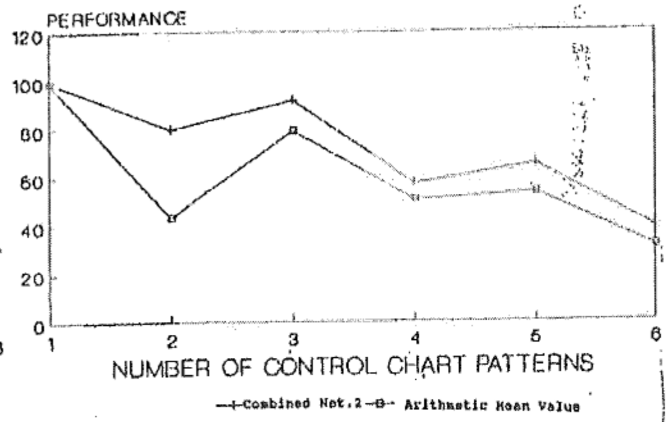

Figure 5.b.

Figure 5.a. Comparison of Performances of Combined Network System Number 1 and Arithmetic Mean Value of The Performances of Five Set of Networks.

Figure 5.b. Comparison of Performances of Combined Network System Number 2 and Arithmetic Mean Value of The Performances of Five Set of Networks.

To study the combined network system, a definition of the performance of learning becomes necessary. This may be defined as:

$$\text{Performance of learning} = \frac{\text{Number of Chart Patterns Recognized}}{\text{Number of Tests}}$$

Figure 6 shows variation of performance of learning with control chart patterns number. It can be seen from figure 6.a that combined network system gives 100% performance for testing the control chart pattern number 1. On the other hand, combined network system gives 80% perforamce of learning for pattern number 2 and 20% for pattern number 4. This result is obtained when testing pattern number 2. However, when testing pattern number 3 with using combined network system, it gives 92% performance of learning for pattern number 3 and 8% performance of learning for pattern number 5. When testing the pattern number 4, the combined network system gives performances of learning 58% for pattern number 4 and 42% for pattern number 2. Similarly, combined network system gives 65% performance of learning for pattern number 5 and 35% performance of learning for pattern number 3, 8% for pattern number 4 and 39% performance of learning for pattern number 6. It can be seen from these figures that combined network system fails to learn 100% for pattern numbers 2, 3, 4, 5 and 6. It confuses with pattern number 2 and 4,

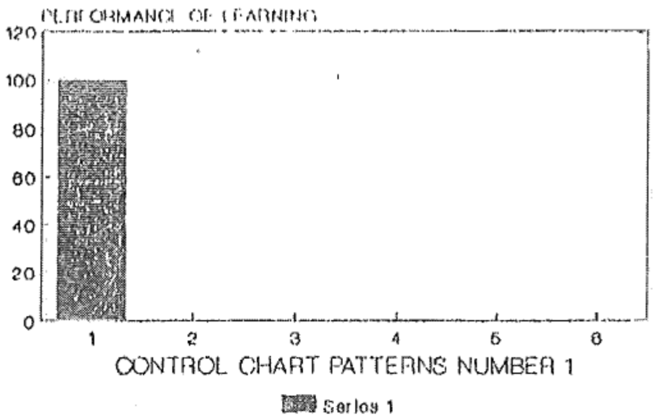

Figure 6.a. Pattern Number 1.

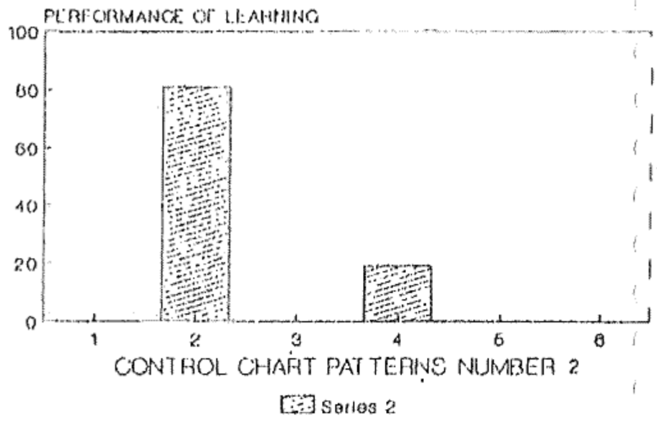

Figure 6.b. Pattern Number 2.

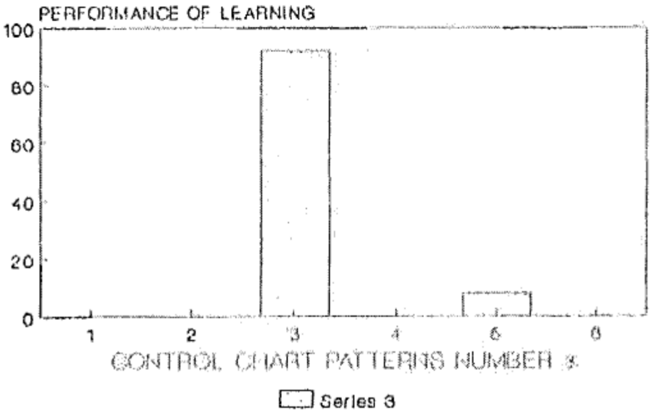

Figure 6.c. Pattern Number 3.

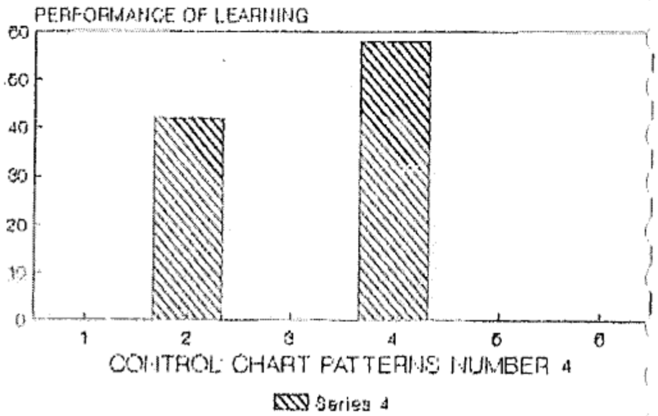

Figure 6.d. Pattern Number 4.

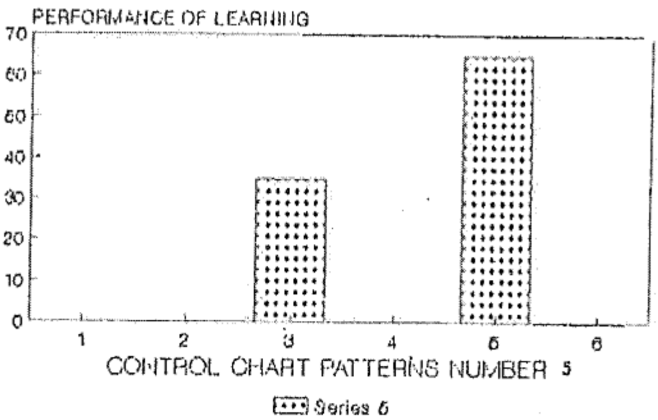

Figure 6.e. Pattern Number 5.

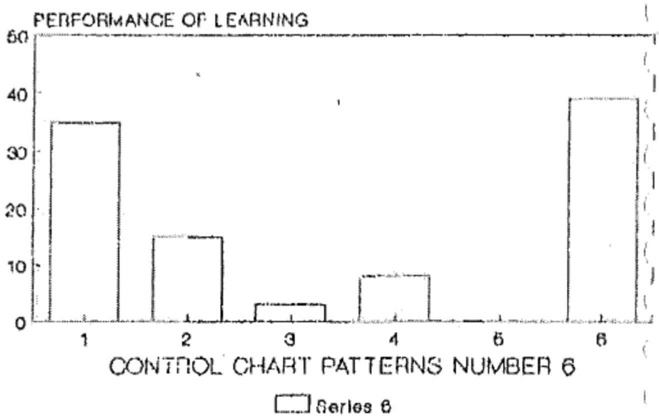

Figure 6.f. Pattern Number 6.

Figure 6. Variation of learning Performances of Combined System With Chart Patterns For Different Pattern Numbers.

and 3 and 5. This may be due to that the amplitudes and frequencies of these patterns are very similar and there may be overlapping of patterns occured. In this case it is very difficult to judge the relevant pattern number (Figure 6.f).

4. CONCLUSIONS

A counterpropagation network fails to recognise and classify the data developed previously for six chart patterns. However, an alternative method for the back-propagation may be developed using a counterpropagation combined network system. In this case the data file should be divided into sub-data files such that each sub-data file contains the data corresponding to six chart patterns. This is necessary, since overlapping of the chart patterns can be avoided. Consequently, counterpropagation networks are trained using different sub-data files and two sets of them are selected due to their better performances. The sixth network is introduced and trained using binary data file and it is added to these network sets which in turn produces a combined network system. The performance of the combined network system is found to be better than the any other individual network's performance. However, the combined network system fails to recognise and classify 100% for some chart patterns. This may be due to the fact that some of the chart patterns have very similar amplitudes and frequencies resulting in overlapping. On the other hand the performance of the combined network system developed in the present study gives improved performances and may be considered as sufficient for the application to the present problem. It is worth to mention here that training time for the combined system is considerably less than backpropagation networks developed for this purpose, i.e about 21000 times trained for the combined network system and over 90000 times trained for the backpropagation for the same sample.

5. REFERENCES

1. Hecht-Nielsen R., Counterpropagation networks, Proc. of IEEE First International Conference on Neural Networks, Part II, (1987) 19-32.

2. Hecht-Nielsen R., Applications of counterpropagation networks, Neural Networks, 2, (1988) 131-140.

3. Roitblat H.L., Moore P.W.B., Nachtigall P.E., Penner R.H. and Au W.W.L., Doplhin echolocation: Identification of returning echoes using a counterpropagation network, Proc. of International Joint Conference on Neural Networks, Part I, (1989) 295-299.

4. Glover D.,A hybrid optical/electronic neurocomputer machine vision inspection system, Proc. of Vision 88 Conference, Society Manufacturing Engineers, Dearborn, MI, (1988).

5. Hertz J., Krogh A and Palmer R.G.,Introduction to the theory of neural computation, Lecture Notes Volume 1, Addison-Wesley Publishing Company, New York, (1991).

6. Lippmann R.P.,Rewiev of neural networks for speech recognition, Neural Computation I, (1989) 1-38.

7. Pham D.T. and Oztemel E.,Using neural networks to recognise patterns of control charts in automated manufacturing, Send to Publisher, (1991).

8. Simpson P.K.,Artificial neural systems: Foundations, paradigms, applications and implementations, Pergamon Press, New York, (1990).