Integrating Neural Network Strategies for Discrimination, Recognition and Clustering

Alexander Iversen, Nicholas K. Taylor, Keith E. Brown
Intelligent Systems Laboratory
Heriot-Watt University
Edinburgh EH14 4AS, UK

***Abstract*— A way of achieving robustness for real-world pattern recognition problems may be through integrating neural network strategies for discrimination, recognition and clustering. Making use of the knowledge that these diverse strategies provide would require us to look at structural design, connectionist versus non-connectionist combination methodologies, and possibly any neurophysiological equivalence.**

I. Problem Statement

The limitations of stand-alone neural network systems have initiated research into multiple neural network systems (multinets), which can be broadly categorised as *ensembles* (diverse neural networks performing the same task) or *modular systems* (similar neural networks performing distinct sub-tasks) [1][1]. The main motivations for introducing these methods are the improvements of accuracy or efficiency respectively [2]. In our view, further progress could be gained by looking at how diverse neural network strategies can be applied to different aspects of pattern recognition problems, unlike the traditional ensemble or modular methods. In particular, we believe we can benefit from integrating the knowledge obtained from discrimination, recognition and cluster-based neural network strategies.

Let us consider an example based on feed-forward neural networks. Consider an MLP used for classification. During training, this neural network is modified so as to create decision boundaries between classes given pattern examples from each class. However, as these decision boundaries are likely to be open, the network cannot reliably detect erroneous inputs and novel patterns, as these would fall into one of the class spaces [3]. This is obviously a problem in *open set problems* such as entry control etc., where we need to identify and give access to known identities and refuse access for unknown identities. However, the problem of open decision boundaries may, in practice, also appear in less obvious cases, i.e., *closed set problems*. For example, we could envisage an MLP diagnosis tool, trained to classify benign and malignant tumours, being presented with a hitherto novel case fundamentally different from any of the training examples. Because of the open decision boundaries, there would always be the chance for a rare malignant case being confidently classified as benign and vice versa. In such, and many other practical situations, recognition based neural networks, e.g. auto-associative MLPs, could be included to detect and isolate irregularities as these networks can create closed decision boundaries or density models based on the training data. Moreover, clustering methods, such as competitive neural networks, may be able to classify the irregular and novel patterns that have been identified.

Unlike many existing multinet systems, whose submodule outputs are often of the same type, the outputs from discrimination, recognition, and cluster based neural networks may be much more diverse and unrelated. The main objective will therefore be to investigate how these neural strategies can be integrated.

II. Current Research

In the last couple of decades, the research into multiple classifier systems, e.g., multi-nets, has increased steadily. For example, recent books have been devoted to the topic [4] [5], and 2005 sees the sixth International Workshop on Multiple Classifier Systems [6].

One of the main themes in multiple classifier systems has been *diversity*. This is simply based on the reason that identical classifiers will make the exact same errors and thus offer nothing in terms of combination. Diversity in modular systems is implicit as each module either process different parts of the system inputs or at least implement different functions. Diversity in ensembles may be realised either through ensemble constituents of different parameter initialisations, types, sizes, topologies etc., or through providing each constituent with different training data. Of the latter, methods that have proved to be popular are Bagging

[1] Note that the terms *ensemble* and *modular* are often used interchangeably in various texts, which can cause some confusion.

[7], which creates diverse training sets by selecting random training samples with replacement, and AdaBoost [8], which involves an adaptive resampling of the training samples as training progresses.

The modules or constituents that make up multiple classifier systems are often of the same type, e.g., only discrimination based classifiers. On the other hand, there are examples of combinations of clustering and discrimination neural networks. In [9] and [10] an ART network is used to decompose the problem by clustering similar input data into clusters and subsequently training discrimination neural networks to handle each subproblem. These methods usually attempt to improve accuracy or efficiency for classification/discrimination but are not aimed at tackling the additional problems of pattern detection, verification or analysis.

Classification and verification have been investigated in [11] and [12] using the MLP for classification/discrimination and auto-associative MLPs for pattern verification. In [11] classification and verification are implemented sequentially to minimise computational cost, whereas in [12] the discrimination and recognition strategies are implemented in parallel so as to exploit the complementary knowledge that the two strategies offer.

III. Key Avenues

We believe that research into multiple classifier systems could benefit by involving further research into the integration of discrimination, recognition and clustering strategies. In particular, this could make such systems more applicable to practical scenarios, as they may offer more reliability when handling sub-optimal or incomplete data. To achieve this goal we believe that the following avenues should be explored:

- *Structural design.* The integration of discrimination, recognition, and cluster based neural strategies may be achieved through a purely sequential structure, e.g., classification followed by verification followed by an analysis of rejected data. However, like diverse networks in a classical ensemble system, the different strategies may also encompass complementary information. For example, a recognition-based neural network may hold information that can give support to discrimination and vice versa. This calls for more complex integration in a parallel structure.

- *Connectionist versus non-connectionist integration approach*. Implicit in this issue is whether to allocate tasks to the neural models a priori or let the allocation become part of the learning process. The choice of integration approach is also influenced by the structural design.

- *Modularity in the brain*. An interesting aspect in the integration of different neural network strategies is to look for any neurophysiological equivalence, i.e., how discrimination, recognition and clustering are encompassed in the brain. As the commonly used neural network models are very simplified abstractions of real neural networks, it might also be possible to create simplified abstractions (connectionist or non-connectionist) based on the interaction of such functions in the brain.

As a final remark, it is important to realise that the strategies proposed in this position statement do not attempt to replace any of the existing models of multiple classifier systems. On the contrary, existing models could well be integrated into such strategies.

References

[1] A. J. C. Sharkey, "Types of multinet systems," In F. Roli and J. Kittler (editors), *Lecture Notes in Computer Science: Proceedings of the Third International Workshop on Multiple Classifier Systems*, vol. 2364, pp. 108-117, Springer-Verlag, 2002.

[2] M. Gori and F. Scarselli, "Are multilayer perceptrons adequate for pattern recognition and verification?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1121-1132, November 1998.

[3] G. Valentini and F. Masulli, "Ensembles of learning machines," In M. Marinaro and R. Tagliaferri (editors), *Lecture Notes in Computer Science: Neural Nets WIRN VIETRI 2002*, vol. 2486, pp 3-19, Springer-Verlag, 2002.

[4] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, New York: John Wiley, 2004

[5] A. J. C. Sharkey, *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems*, London: Springer-Verlag, 1999.

[6] International Workshop on Multiple Classifier Systems: http://www.diee.unica.it/mcs/.

[7] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, no. 2, pp. 123-140, August 1996.

[8] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," *Proceedings of the Thirteenth International Conference on Machine Learning,* Bari, Italy, 1996, pp. 148-156.

[9] G. Auda and M. Kamel, "CMNN: Cooperative modular neural networks for pattern recognition," *Pattern Recognition Letters*, vol. 18, pp. 1391-1398, November 1997.

[10] W. Tsai, H. Tai and A. C. Reynolds, "An ART2-BP supervised neural net," *Proceedings of the IEEE International Conference on Neural Networks*, San Diego, USA, 1994, vol. 5, pp. 3289-3294.

[11] A. Iversen, N. K. Taylor, K. E. Brown and J. Kårstad, "Hybrid neural network classifying recognisable and detecting unrecognisable radio signals," *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications*, Innsbruck, Austria, 2005, pp. 814-819.

[12] A. Iversen, N. K. Taylor and K. E. Brown, "Classification and verification through the combination of the multi-layer perceptron and auto-association neural networks," To appear in *Proceedings of the International Joint Conference on Neural Networks*, Montréal, Canada, 2005.