Experiences with Neural Networks as a Diagnostic Tool in Medical Image Processing

Georg Dorffner
Dept. of Medical Cybernetics and Artificial Intelligence
University of Vienna

Erich Prem, Markus Mackinger, Stefan Kundrat, Paolo Petta
Austrian Research Institue for Artificial Intelligence

Gerold Porenta, Heinz Sochor
Department of Cardiology, Clinic of Internal Medicine 2
University of Vienna Medical School

1 Introduction

Through recent years artificial neural networks have proven to be a useful technique in the interpretation of high-dimensional data such as images. However, an adequate application of neural networks is often plagued by a lack of systematic methodology. It is one of the goals of the EC funded ESPRIT-II project NEUFODI (Neural Networks for Forecasting and Diagnosis Applications) to study and develop techniques for using artificial neural networks as tools for diagnosis.

In this paper we report about part of this larger project, namely about experiences with applying neural networks to the interpretation of planar thallium-201 scintigrams [4] for the assessment of coronary artery disease. This application should serve as an example of how neural networks can be successfully applied in the area of medical image processing. In this realm we will discuss several aspects about the practical use of this widely used technique.

2 Neural Networks for Diagnosis

Artificial neural networks [2] - or neural networks, for short - are a popular approach to information processing in computer science and artificial intelligence. Their name is derived from the fact that some basic ideas of processing were adapted from our knowledge about the function of groups of nerve cells in the brain. The main features are massively parallel processing in a large group of relatively simple but highly interconnected processors (or 'units'), and self-organization or adaptation through so-called learning algorithms that change the connectivity between the units. Neural networks can be trained by samples from a specific domain and thus obtain their "knowledge" about appropriate processing through extracting important information from those samples. Among others, supervised and unsupervised learning schemes can be distinguished depending on the role and specificity of a "teacher" (or feedback). In the domain of diagnosis the most typical form of neural networks consists of several layers of units, each unit in a layer being connected to each unit in the next one. Such networks - among others known as multilayer perceptrons or feedforward associative networks - can be trained to implement mappings from input data to output activations, most of the time representing one of several diagnoses.

3 The Assessment of Coronary Artery Disease

Since 1988, in a cooperation between the Austrian Research Institute for Artificial Intelligence and the Dept. of Cardiology at the University of Vienna, we have been investigating the possible use of neural networks in the assessment of coronary artery disease (CAD) [5,6]. The reference method for this is coronary angiography, which has the disadvantage of being an invasive procedure. Among other non-invasive tests myocardial scintigraphy , which relies on expert readings, has proven useful in the detection of CAD [4]. Diagnoses related to CAD concern the presence of disease (yes/no), the localization (vascular territory) or the extent of disease (single or multivessel).

The very nature of the interpretation of scintigrams lends itself nicely to an automated solution employing neural networks. In particular, data from thallium-201 scintigraphy was chosen as the source for inputs to a three-layer perceptron, while the output was designed to correspond to one of the three aforementioned ways of classification. The process of diagnosis in this sense is a one-shot mapping from the input data to the output. This means that the involved knowledge (in the sense of artificial intelligence) or structures have to be considered as low-level, as opposed to what many medical expert systems have to deal with [1]. At the same time this technique appears easily generalizable to other similar problems of classifying medical images.

Planar thallium-201 scintigraphy generates images of the myocardial isotope uptake after stress and during resting conditions in three centered views of the heart (i.e. the three standard projections anterior, LAO 45 degrees, LAO 70 degrees).

Image processing in its most general sense implies the treatment of very large sets of data. In our case the original scintigrams are stored on computer file as 64 x 64 pixel matrices. The complete encoding of such an image would require a huge number of input units, which, for several reasons, makes it hard to be handled by common computers. First, large input patterns imply extreme storage requirements, which in many cases could however be solved. Secondly, the number of input units increases the number of connections and thus the training time. Thirdly, and most severely, training networks with large numbers of connections would require a similarly large number of training cases, considering the fact that the network has to achieve optimization of a process involving as many degrees of freedom as connections. In our case, around 150 patient records were available, far too few to match the number of connections in a full-fletched network taking entire images as input. For these reasons, a method of reducing the size of a single input pattern was desired from the beginning. The method we employed is based on circumferential profile analysis and generates normalized count rates in five anatomical segments per planar view. In addition, segmented washout rates were obtained. All results, briefly reported below, were achieved using this reduced input representation of 45 values. This shows that even though images without preprocessing easily grow out of hand, in many cases efficient solutions can be achieved from applying compression techniques preserving the relevant information.

For most of the experiments 159 data sets were available, between 10 and 30 of which were used during training (leaving the rest as a test set for evaluating the generalization performance). Alternatively, the diagnosis from expert readings of the scintgrams or from angiography were used as target output. The network consisted of 45 input units (corresponding to the values as described above), 5 to 30 hidden units, and 1 to 3 output units (depending on the type of diagnosis). Different variations of the backpropagation learning rule [7] were applied. The goal for the initial studies was to find the network with the best performance on the test set. The test set was varied in order to achieve this goal, while keeping it as small as possible to increase the confidence level of the results.

The best results were obtained for predicting the presence of disease (ranging from 87 to 91 % correct for the test set). Sensitivity was higher for the network trained on the expert readings, while the network trained on angiography output had a larger specificity. Results on location (78 - 81 %) and extent (78 - 82 %) were lower but still highly useful. Here the network trained on the expert was clearly better as compared to angiography (67 %). These results are better, while using fewer inputs, than comparable approaches [3].

4 Using Knowledge to Improve Performance

The common view of neural networks and rule-based approaches to diagnosis tasks is that the former always acquire their knowledge through training, while in the latter case knowledge is pre-programmed as rules. Thus, each method has its type of applications where they appear most appropriate, in the case of neural networks applications where no or little explicit knowledge exists or is known. As a result, neural networks are usually trained with random inital configurations, the only source of knowledge being the training examples. In many real world cases such as the one presented here, however, partial symbolic knowledge on how to solve the problem exists beforehand and is already formalized. It seems to be a waste of effort not to make use of this knowledge. The following are possible reasons why finding a way of inserting explicit knowledge into a neural network can be desirable.

- Learning from scratch (*tabula rasa*) is a very costly process. When starting with randomly distributed weights, the training examples have to be presented to the network many times before any success becomes visible.

- The number of training samples needed to train an "empty" network can be very large. In medical applications often only a restricted number of patient data is available, partially due to complex and costly methods of investigations, partially due to low numbers of existing cases.

- Not all the necessary knowledge might actually be contained in the training samples themselves. A data set might not contain the most typical cases, or a diagnosis might best be based on information not immediately evident in the data (such as dependencies on the patient's sex).

In all these cases the use of existing knowledge can help overcome the limitations of the brute force neural network method. Thus we have developed a specific technique, called ***concept and rule support***, of pretraining a neural network on rules or supportive concepts [6]. An example for a rule in interpreting thallium-201 scintigrams is the following.

> *If a segmental uptake value is below 60 or the washout rate is negative then classify the case as pathological.*

All rules of this kind, which were derived from interviewing an expert, cover only about 75 % of the cases. Thus they by themselves are too weak to justify a rule-based expert system approach. However, not making use of this body of knowledge would seem to be a waste of resources. The results show that applying the method of concept and rule support consistently improved the generalization performance of the trained network, compared to reference tasks with the same conditions but without support.

5 Implementations

The described neural networks have been implemented by the simulation tool box VieNet2 (or its earlier version VIE-NET), developed at the Austrian Research Institute for Artificial Intelligence. This tool box runs on both workstations and IBM compatibles and is written in C, including a graphical interface. This permitted rapid prototyping and gradual refinement of the network architectures, as well as easily portable final implementations. The results on including symbolic knowledge were achieved by using the commercial neural network tool Brainmaker (TM) [8]. In addition, an easily usable program including a datalink interface between an Apex-Elscint gamma camera and a MacIntosh computer was written to make the final system widely available. A cross-evaluation study at several clinics is planned. Such a validation study would be the final step in establishing this relatively novel technique as a viable diagnostic tool.

6 Experiences Gained

The experiences in diagnosing thallium-201 scintigrams have demonstrated that neural networks can effectively be used as a tool in medical decision making. In summary, the following aspects of interest have suggested themselves during the extensive testing of this method.

- The described application has demonstrated how otherwise costly image processing tasks can effectively be reduced to manageable network size when employing a technique of reducing data without reducing relevant information. Although this cannot automatically be generalized to other applications, it nevertheless promises similar successes in a variety of different imaging tasks.

- The method of incorporating knowledge has shown a way of combining neural networks with more traditional approaches of automatic medical diagnosis.
- The acceptability of this method by persons who actually have to apply it in clinical routine appears to be higher, as compared to more sophisticated expert system in artificial intelligence. The main reason for this is its lower-level nature of being a one-shot system. A neural network of this kind, when built into an easy-to-understand user interface, is simple to apply and does not require large online training.

  Among experts, on the other hand, the acceptability tends to be lower due to the well-known fact that neural networks are not really capable of "explaining" their line of reasoning. Further research into hybrid symbolic/neural network approaches might help overcome this problem.
- As in traditional artificial intelligence in medicine, or even more so, diagnostic tools of the described kind cannot be expected to somehow replace human experts. Instead, they are expected to be of high value as one important link in the chain of clinical decision making. Further possible applications of neural networks as a diagnostic tool are in teaching students, training younger physicians, and in the preliminary assessment of large volume data.

In conclusion, we expect to see neural networks as an established technique for both automatic medical imaging and diagnosis in the near future. By making use of international projects, such as the initally mentioned NEUFODI, European medical research has the chance of staying apace with this important development.

7 Acknowledgments

The first year of this research leading to the basic results and a ready-to-use implementation was sponsored by Informatica Ges.m.b.H. Austria. More recently this work has been part of the EC-funded ESPRIT-II project 5433 NEUFODI ("Neural Networks for Forecasting and Diagnosis Applications"). Partners in this project are the Austrian Research Institute for Artificial Intelligence, the Babbage Institue for Knowledge and Information Technology (Belgium), Lyonnaise des Eaux-Dumez (France), Labein Reserach Labs (Spain), and Elorduy y Sancho (Spain). The Austrian contribution was supported by the Austrian Industrial Research Promotion Fund as grant no. 2/282.

References

[1] B.G.Buchanan, E.H.Shortliffe (eds.)*Rule-Based Expert Systems - The MYCIN Experiments of the Stanford Programming Project*, Addison-Wesley, Reading, MA, 1984.

[2] R.Hecht-Nielsen *Neurocomputing*, Addison-Wesley, Reading, MA, 1990.

[3] H.Fujita et al. *Application of Artificial Neural Network to Computer-Aided Diagnosis of Coronary Artery Disease in Myocardial SPECT Bull's-eye Images*, Journal of Nuclear Medicine, 33(2)272-276, 1992.

[4] Myron C.Gerson *Thallium-201 Two-dimensional Imaging.* In:M.J.Gelfand, S.R.Thomas (Eds.), Effective Use of Computers in Nuclear Medicine, Mc Graw-Hill, New York, 1988, pp 75-108.

[5] G.Porenta, G.Dorffner, J.Schedlmayer, H.Sochor *Parallel Distributed Processing as a Decision Support Approach in the Analysis of Thallium-201 Scinitgrams.* In: Proc.Computers in Cardiology 1988, Washington D.C., IEEE.

[6] E.Prem, M.Mackinger, G.Dorffner, G.Porenta, H.Sochor *Concept Support as a Method for Programming Neural Networks with Symbolic Knowledge.* to appear in: Proceedings of the German Workshop on AI (GWAI), Springer Heidelberg, 1993.

[7] D.E.Rumelhart, G.E.Hinton, R.J.Williams *Learning Internal Representations by Error Propagation*, in Rumelhart D.E., McClelland J.L.(eds.), Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Vol 1: Foundations, MIT Press, Cambridge, MA, 1986.

[8] - *Brainmaker - User's Guide and Reference Manual*, California Scientific Software, 1990.