Brain Signal Analysis

Jeng-Ren Duann, Tzyy-Ping Jung, Scott Makeig

Institute for Neural Computation, University of California, San Diego CA
Computational Neurobiology Lab, The Salk Institute, La Jolla CA

RUNNING HEAD: Brain signal analysis

Artificial neural networks (ANNs) have now been applied to a wide variety of real-world problems in many fields of application. The attractive and flexible characteristics of ANNs, such as their parallel operation, learning by example, associative memory, multifactorial optimization and extensibility, make them well suited to the analysis of biological and medical signals. In this study, we review applications of ANNs to brain signal analysis, for instance, for analysis of the electroencephalogram (EEG) and magnetoencephalogram (MEG), or electromyogram (EMG), and as applied to computed tomographic (CT) images and magnetic resonance (MR) brain images, and to series of functional MR brain images (i.e. fMRI).

1. INTRODUCTION

Artificial neural networks (ANNs) are computational framework inspired by our expanding knowledge of the activity of networks of biological neurons in the brain. ANNs cannot hope to reproduce all the still not well-understood complexities of actual brain networks. Rather, most ANNs are implemented as sets of nonlinear summing elements interconnected by weighted links, forming a highly simplified model of brain connectivity. The basic operation of such artificial neurons is to pass a weighted sum of their inputs through a nonlinear hard-limiting or soft "squashing" function. To form an ANN, these basic calculating elements (artificial neurons) are most often arranged in interconnected layers. Some neurons, usually those in the layer furthest from the input, are designated as output neurons. The initial weight values of the interconnections are usually assigned randomly.

The operation of most ANNs proceeds in two stages. Rules used in the first stage, training (or learning), can be categorized as supervised, unsupervised, or reinforced. During training, the weight values for each interconnection in the network are adjusted either to minimize the error between desired and computed outputs (supervised learning) else to maximize differences (or to minimize similarities) between the output categories (unsupervised or competitive learning). In reinforced learning, an input-output mapping is learned during continued interaction with the environment so as to maximize a scalar index of performance (Haykin, 1999). The second stage is recall, in which the ANN generates output for the problem the ANN is designed to solve, based on new input data without (or sometimes with) further training signals.

Because of their multifactorial character, ANNs have proven suitable for practical use in many medical applications. Since most medical signals of interest are usually not produced by variations in a single variable or factor, many medical problems, particularly those involving decision-making, must involve a multifactorial decision process. In these cases, changing one variable at a time to find the best solution may never reach the desired objective (Dayhoff and DeLeo, 2001), whereas multifactorial ANN approaches may be more successful. In this chapter, we review recent applications of ANNs to brain signal processing, organized according to the nature of brain signals to be analyzed and the role that ANNs play in the applications.

2. ROLES OF ANNS IN BRAIN SIGNAL PROCESS

To date, ANNs have been applied to brain data for the following purposes:

a. *Feature extraction, classification, and pattern recognition:* ANNs here serve mainly as non-linear classifiers. The inputs are preprocessed so as to form a feature space. ANNs are used to categorize the collected data into distinct classes. In other cases, inputs are not subjected to preprocessing but are given directly to an ANN to extract features of interest from the data.

b. Adaptive filtering and control: ANNs here operate within closed loop systems to process changing inputs, adapting their weights "on the fly" to filter out unwanted parts of the input (adaptive filtering), or mapping their outputs to parameters used in online control (adaptive control).

c. *Linear or nonlinear mapping:* Here ANNs are used to transform inputs to outputs of a desired form. For example, an ANN might remap its rectangular input data coordinates to circular or more general coordinate systems.

d. *Modeling:* ANNs can be thought of as function generators that generate an output data series based on a learned function or data model. ANNs with two layers of trainable weights have been proven capable of approximating any nonlinear function.

e. *Signal separation and deconvolution:* These ANNs separate their input signals into the weighted sum or convolution of a number of underlying sources using assumptions about the nature of the sources or of their interrelationships (e.g., their independence).

f. *Texture analysis and image segmentation:* Image texture analysis is becoming increasingly important in image segmentation, recognition and understanding. ANNs are being used to learn spatial or spatial-frequency texture features and, accordingly, to categorize images or to separate an image into subimages (image segmentation).

g. *Edge detection:* In an image, an edge or boundary between two objects can be mapped to a dark band between two lighter areas (objects). By using the properties of intensity discontinuity, ANNs can be trained to “recognize” these dark bands as edges, or can learn to "draw" such edges based on contrast and other information.

3. APPLICATION AREAS

In this section, we illustrate applications of ANNs to brain signals through some examples involving neurobiological time series and brain images. Neurobiological signals of clinical interest recorded noninvasively from humans include EEG, MEG, and EMG data. Research in brain imaging includes the analysis of structural brain images, mainly focused on the extraction of 3-dimensional structural information, from various kinds of brain images (e.g., magnetic resonance images, MRI), as well as analysis of functional brain imaging series that mainly reveal changes in the brain state during cognitive tasks using medical imaging techniques (e.g., fMRI and positron emission tomography or PET). These examples, however, by no means cover all the publications in the field, whose number is growing rapidly.

Neurobiological signals

- **Electroencephalogram and Magnetoencephalogram**

The electroencephalogram (EEG) is a non-invasive measure of brain electrical activity recorded as changes in the potential difference between two points on the scalp. The magnetoencephalogram (MEG) is its magnetic counterpart. In accordance with the assumption that the ongoing EEG can be alternated correspondingly by stimulus or event to form the event-related potential (ERP) or the evoked potential (EP), these changes, though tiny, can be recorded through the scalp. It is possible for researchers to apply pattern recognition algorithms to search for the differences in brain status while the brain is performing different tasks. Thus, Peters and colleagues (2001) applied an autoregressive (AR) model to four-channel EEG potentials to obtain features that were used to train an ANN using a backpropagation algorithm to differentiate the subject's intention to move the left or right index finger or right foot. They suggested the framework might be useful for designing a direct brain-computer interface. In the study of Zhang et al. (2001), ANNs were trained to determine the stage of anesthesia based on features extracted from the middle-latency auditory evoked potential (MLAEP) plus other physiological parameters. By combining power spectral estimation, principal component analysis and ANNs, Jung et al. (1997) demonstrated that continuous, accurate, noninvasive, and near real-time estimation of an operator's global level of alertness is feasible using EEG measures recorded from as few as two scalp sites. Results of their ANN-based estimation compare favorably to those using a linear regression model applied to the same PCA-reduced EEG power spectral data.

As a linear mapping device, Sun and Sclabassi (2000) employed an ANN to transform the EEG topography obtained from a forward solution in a simple spherical model to a more realistic spheroidal model whose forward solution was difficult to compute directly. Here, a backpropagation learning algorithm was used to train an ANN to convert spatial locations between spherical and spheroid models. Instead of computing the infinite sums of the Legendre functions required in the asymmetric spheriodal model, the calculations were carried out in the spherical model and then converted by the ANN to the more realistic model for display and evaluation.

Recently, ANNs have made an important impact on the analysis of EEG and MEG by separating the problem of EEG or MEG source identification from that of source localization, a mathematically underdetermined problem -- any scalp potential distribution can be produced by a limitless number of potential distributions within the head. Because of volume conduction through cerebrospinal fluid, skull and scalp, EEG and MEG data collected from any point on the scalp may include activity arising in multiple locally synchronous but relatively independent neural processes within a large brain volume. This has made it difficult to relate EEG measurements to underlying brain processes and to localize the sources of EEG and MEG signals. Progress has been made by several groups in separating and identifying the distinct brain sources from their mixtures in scalp EEG or MEG recordings assuming only their temporal independence and spatial stationarity (Makeig et al., 1997; Jung et al., 2001), using a class of independent component analysis (ICA) or blind source separation (BSS) algorithms.

- **Muscle and Movement Signals**

From recordings of muscle stretching (mainly, the electromyogram or EMG), it is possible to predict the intent of subjects to perform actions such as hand or finger movements, or to judge the disability of a specific bundle of muscle cells. For example, Khalil and Duchene used wavelet coefficients obtained from uterine electromyography to train ANNs (Khalil and Duchene, 2000) to separate the inputs into four labeled categories: uterine contractions, fetal movements, Alvarez waves, and long-duration low-frequency band (LDBF) waves. They reported that the system was useful for maintaining preterm births. On the other hand, Stites and Abbas (2000) used an ANN as a pattern shaper to refine the output patterns of a functional neuromuscular stimulation system (FNS) that served as a pattern generator of control signals for cyclic movements to help the paraplegic patient stand using FNS.

Brain Images

- **Structural Images**

In structural brain image analysis, ANNs may play roles in image segmentation, image labeling and/or edge detection. Image segmentation is the first, and probably the most important step in digital image processing. Segmentation may be a labeling problem in which the goal is to assign, to each voxel in a gray-level image, a unique label that represents it’s belonging to an anatomical structure. The results of image segmentation can be used for the image understanding and recognition, three-dimensional reconstruction, visualization, and for measurements including brain volume changes in developmental brain diseases such as Alzheimer's disease and autism. The rapid pace of development of medical imaging devices such as magnetic resonance imaging (MRI) and computerized tomography (CT), allows to better understanding of anatomical brain structure without, prior to, or even during neurosurgery. However, results are highly dependent upon the quality of the image segmentation processes.

Here, we give some examples using ANNs in image segmentation: Dawant et al. (1991) presented a backpropagation (BP) neural network approach to the automatic characterization of brain tissues from multi-modal MR images. The ability of a three-layer BP neural network to perform segmentation based on a set of MR images (T1-weighted, T2-weight and proton density weighted) acquired from a patient was studied. The results were compared to those obtained using a Maximum Likelihood Classifier. They showed there was no significant difference in the results obtained by both methods, though BP neural network gave cleaner segmentation images. By using the same analysis strategy, Reddick et al. (1997) first trained a self-organizing map (SOM) on multi-modal MR brain images to efficiently extract and convert the 3-D inputs (from T1-, T2- and PD-weighted images) into a feature space and utilized a BP neural network to separate them into classes of white matter, gray matter, and cerebral spinal fluid (CSF). Their work demonstrated high intraclass correlation between the automated segmentation and classification of tissues and standard radiologist identification as well as high intrasubject reproducibility.

- **Functional Images**

Nowadays, not only does medical imaging device provide impressive spatial resolution and details of the fine structure of the human brain, it is also able to reveal changes in brain status while awake subjects perform a task or even daydream by measuring ongoing metabolic changes including cerebral blood flow (CBF), cerebral blood volume (CBV) (by Positron Emission Tomography, PET), and blood oxygenation level-dependent (BOLD) signal levels (by functional MR imaging, fMRI). We will give some examples mainly from fMRI analysis.

Functional brain imaging emerged in the early 90's based on the observation that increase in local neuronal activity are followed by local changes in oxygen concentration. Changing the amount of oxygen carried by hemoglobin changes the degree to which hemoglobin disturbs a magnetic field is able to demonstrate that in vivo changes blood oxygenation could be detected by MRI (Ogawa et al., 1992). The subsequent changes in the MRI signal became known as the blood-oxygenation- level-dependent or BOLD signal. This technique was soon applied to normal humans during functional brain activation, by cognitive task performance, giving birth to the rapid growing field of functional magnetic resonance imaging.

Theoretically, the fMRI BOLD signal from a given brain voxel can be interpreted as a linear combination of different sources with distinguishable time courses and spatial distributions, including use-dependent hemodynamic changes, blood or central spinal fluid flows, plus subject movement and machine artifacts. Recently, ANNs (especially independent component analysis, ICA), applied to fMRI data, have proven to be a powerful method for detecting and separating task-related activations with either known or unanticipated time courses (McKeown et al., 1998) that could not be detected by standard hypothesis-driven analyses. Duann et al. (2002) have given further details of applying ICA to fMRI BOLD signal showing that the hemodynamic response to even widely spaced stimulus presentations may be trial, site, stimulus and subject dependent. Thus, the standard regression–based method of applying a fixed hemodynamic response model to find stimulus- or task-related BOLD activations needs to be reconsidered.

4. DISCUSSION

Uses of ANNs as classifiers currently dominates their applications to the field of brain signal analysis. This includes classification of brain or related signals as exhibiting normal or abnormal features or processes. Not surprisingly, all published studies report promising results.

If the measurements can be modeled as an additive mixture of different sources, including task-related signals and artifacts, applying blind source separation (BSS) prior to the further processing, visualization, or interpretation may better reveal the underlying physical phenomena such as different brain processes) which in the raw data could be contaminated or overwhelmed by other processes of no interest.

A survey of relevant papers shows that the most popular architecture for artificial neural network used is the multilayer perceptron (MLP). The MLP architecture is both simple and straightforward to implement and use. In MLPs, information flows in one direction except during training, when error terms are back-propagated. Backpropagation updates network weights in a supervised manner. Although it cannot guarantee a globally minimal solution, backpropagation at least arrives at a local minimum through gradient descent. Various techniques have been derived to attempt to avoid overfitting to a local minimum. Once the network weights have been learned and fixed, feedforward networks can be implemented in hardware and made to run in real-time. All these characteristics make the backpropagation algorithm most popular in biomedical applications.

In some applications, target outputs may not be available or may be too expensive to acquire. In these cases, unsupervised learning algorithms may be used. Among unsupervised learning algorithms, self-organizing maps (SOMs) are the most popular for biomedical applications. During training, SOMs attempt to assign their input patterns to different output regions. Often

SOMs may converge after only few learning cycles.

APPLICATION ISSUES

Although most published papers have concluded that ANNs are appropriate for their domain of interest, many issues still have to be resolved before ANNs may be claimed to be the general method of choice. Unfortunately, most published studies have not gone beyond demonstrating application to a very limited amount of data. As with any type of method, ANNs have their limitations that should be carefully considered:

- Every study should provide a rationale for the data chosen as input. For example, ANN-based computer-aided-diagnosis (CAD) systems may give misleading results if the ANNs are not given adequately representative features and sufficient naturally occurring data variations in their training data. Using ANNs, any input may yield some sort of output, correct and useful or not (“garbage in, garbage out”). Therefore, keys for success of ANN applications are not only to pick an appropriate architecture or learning algorithm, but also to choose the right data and data features to train the network.
- Although methods of applying ANNs to biomedical signals have already shown great promise, great care must be taken to examine the results obtained. The issue of trust in the outputs of ANNs always deserves informed as well as statistical consideration. Since medical diagnosis is nearly always a multifactorial and multidisciplinary problem, medical experts should always evaluate network outputs in light of other direct or indirect convergent evidence before making final decisions affecting the health of patients.
- Before practical implementation is planned, ANN methods should be compared to more direct ways of obtaining the same answers, as these might sometimes prove more accurate or cost-effective.

MODEL MINING

Since the first wave of popularization of backpropagation networks, nearly two decades ago, an ever-growing number and variety of ANN models have been devised to tackle an ever-widening variety of problems. The overall insight that ANNs both embody and exemplify is perhaps that our human intelligence is multifactorial and highly adaptable to using whatever forms of information are available to us. In this spirit, we suggest that researchers always attempt to interpret the physiological meaning both of the features of their input data and of the data models that their trained ANNs represent. Too often ANNs have been treated like “black boxes.” We believe it is time to open the black boxes and interpret what is happening inside them. Such interpretations might even give new insights into the nature of the biomedical signals, or suggest new or more efficient ways to look at the input data. It is also possible that the ANN models and methods might suggest more efficient methods to collect input data. Such 'model mining' might even prove to be the most rewarding result of applying ANNs. Researchers who simply recount classification accuracy may ignore nuggets of novel information about brain processes hidden in the ANN models that they and the data have jointly constructed.

REFERENCES

Dawant B. M., Ozkan, M., Zijdenbos, A. and Margolin, R., 1991, A computer environment for 2D and 3D quantitation of MR images using neural networks. Mag. Reson. Imaging, 20: 64-65.

* Dayhoff, J. E. and DeLeo, J. M., 2001, Artificial neural networks: opening the black box. Cancer, 91: 1615-1635.

* Duann, J. R., Jung, T. P., Kuo, W. J., Yeh, T. C., Makeig, S., Hsieh, J. C. and Sejnowski, T. J., 2002, Single-trial variability in event-related BOLD signals. NeuroImage, 15: 823-835.

* Haykin, S., 1999, Neural Network: A Comprehensive Foundation, New Jersey: Prentice Hall International Inc.

Jung, T.-P., Makeig, S., Stensmo, M., and Sejnowski, T. J., 1997, Estimating alertness from the EEG power spectrum, IEEE Trans Biomedical Engineering, 44: 60-69.

* Jung, T.-P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T.-W. and Sejnowski, T. J., 2001, Imaging brain dynamics using independent component analysis, Proc. of the IEEE, 89: 1107-1122.

Khalil, M. and Duchene, J., 2000, Uterine EMG analysis, a dynamic approach for change detection and classification, IEEE Trans Biomedical Engineering, 47: 748-756.

* Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., and Sejnowski, T. J., 1997, Blind separation of auditory event-related brain responses into independent components, Proc. Natl. Acad. Sci. USA., 94: 10979-10984.

* McKeown, M. J., Jung, T.-P., Makeig, S., Brown, G. G., Kindermann, S. S., Lee, T.-W. and Sejnowski, T. J., 1998, Spatially independent activity patterns in functional MRI data during the stroop color-naming task, Proc. Natl. Acad. Sci. USA., 95: 803-10.

* Ogawa, S., Tank, D., Menon, R., Ellermann, J., Kim, S., Merkle, H., and Ugurbil, K., 1992,

Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging, Proc. Natl. Acad. Sci. USA., 89: 5951-5959.

Peters, B. O., Pfurtscheller, G., and Flyvbjerg, H., 2001, Automatic differentiation of multichannel EEG signals, IEEE Trans Biomedical Engineering, 48: 111-116.

Reddick, W. E., Glass, J. O., Cook, E. N., Elkin, T. D. and Deaton R. J., 1997, Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks, IEEE Trans Medical Imaging, 16: 911-918.

* Stites, E. C. and Abbas, J. J., 2000, Sensitivity and versatility of an adaptive system for controlling cyclic movements using functional neuromuscular stimulation, IEEE Trans Biomedical Engineering, 47: 1287-1292.

* Sun, M. and Sclabassi, R. J., 2000, The forward EEG solution can be computed using artificial neural networks, IEEE Trans Biomedical Engineering, 47: 1044-1050.

Zhang, X.-S., Roy, R. J., Schwender, D., and Daunderer, M., 2001, Discrimination of anesthetic states using mid-latency auditory evoked potential and artificial neural networks, Ann. Biomedical Engineering, 29: 446-453.