

Use of artificial neural network in diagnostic pathology

Javed Iqbal Kazi, Muhammed Mubarak

Department of Histopathology, Sind Institute of Urology and Transplantation, Civil Hospital, Karachi.

Histopathology is almost entirely based on subjective human interpretation of visual images. The dominant role of human interpretation is explained by the complexity of the images seen down the microscope and the very efficient processing of this information by the human brain. Evolution has produced *Homo sapiens* with high resolution vision and a large biological neural network to interpret those data. Pathologists have been trying for some time to combine their human skills in histopathological diagnosis with the advantages offered by computer systems. Expert systems are computer programs designed to process knowledge and reach the diagnostic decision in the same way as human experts.¹⁻⁵ Methods which allow diagnostic knowledge to be represented with in a computer program include rules, artificial neural networks (ANNs) and inference networks.¹ ANNs can compete with human performance in areas where constant vigilance or objective weighting of multiple probabilities are required.

What is an artificial neural network?

An artificial neural network (ANN) is a nonlinear, computational, mathematical model for information processing, with architectures inspired by neuronal organizational biology. Historically, the first use of the concept of ANNs is attributed to Frank Rosenblatt who, in 1958, used contemporary knowledge of neurophysiology to create a mathematical model to simulate information processing in human neurons.⁵ This early design was called a "perceptron" and remains the basis for some of today's advanced ANN systems. Figure 1 is a diagram of the most commonly used architecture of artificial neural network, the multilayer perceptron (MLP). The network is constructed of

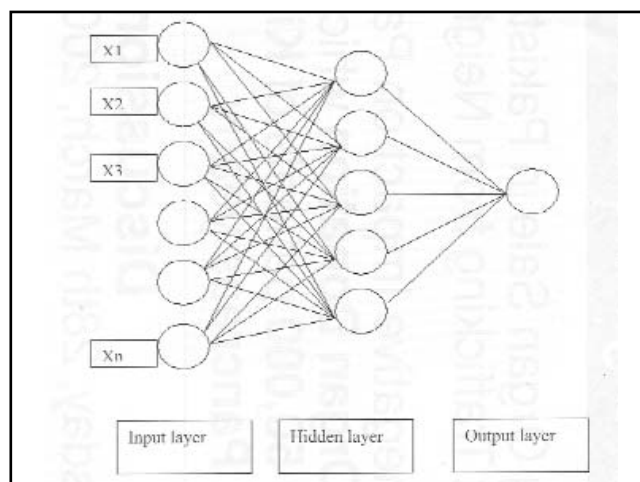


Figure 1. An artificial neural network with the multilayer perceptron structure. The network shown has an input layer of 'n' neurons, $X_1, X_2, \dots, X_n$, one for each input to the network, one hidden layer, and an output layer of one neuron.

discrete artificial neurons or nodes (shown as circles) which are interconnected by routes or links that correspond to the axons in biological neurons. Like biological neurons, an artificial neuron summates the signal arriving from incoming connections. If this reaches a certain threshold, a signal is fired down the outgoing axons. The input data are presented to the layer of input neurons or evidence node with one input neuron required for each item of input data. The prediction of the neural network is given by the layer of output neurons or decision node. If an artificial neural network is constructed that consists solely of an input layer connected directly to an output neuron, then this network will function only as a linear regression classifier. If a hidden layer is interposed between input and output layers then the network will be able to perform more complex classification tasks.^{2,4,5}

Network training and initialization

The important aspects of developing a neural network system are the general principles of the training and testing cycle.^{2,5} A newly constructed artificial neural network will have random weighting (i.e. initial probability) applied to all the connections within the network and will give random predictions. For the network to have any useful function it must be trained or initialized with input data (such as histological features) which have a known outcome verified by a gold standard method and the weights of the connections changed to give correct predictions by a feedback mechanism. The particular method of minimizing errors will vary among different network architectures. In MLPs the most common method of learning is by back propagation in which the difference between the net's prediction and the true outcome is referred back through the network and minimized by changing the weight of connections (i.e. conditioned probability). The initialization process will thus provide a conditioned probability matrix (CPM) for the program. To define the CPM for each possible result for each feature one must answer the question: if this feature is present at this intensity, what is the probability that the specified outcome (eg. acute rejection) is actually present? The quality of these data obviously influences the quality of the output. It can successfully be provided merely as expert opinion, but results are likely to be better if they are based on actual observations. The probability needs to be adjusted slightly to allow for the possibility of rare events, which may not be represented in representative series of training biopsies. A probability of 0 or 100% at any point in the CPM would lead the network to give a diagnosis with a spurious absolute certainty and invalidate the consideration of any other features. Therefore, the figures need to be manually smoothed to remove such extremes. The resulting figures for CPM of all the variables are entered into the computer program.

The size of training set will vary with the number of different input and output variables, the number of neurons in hidden layers and the ranges that the variables have in the problem domain. A general rule of thumb is to have considerably more cases in training set than the total number of connections within the network.^{1,2,5}

Applications in Medicine

The artificial neural networks were originally developed by psychologists who hoped that simulations of human cerebral function would allow the investigation of diseases such as schizophrenia. That expectation has been largely unrealized and ANNs are used now as sophisticated statistical classifiers and image or signal processing devices.¹⁻³ Given the strength of ANNs in performing pattern recognition tasks, early clinical applications were in diagnostic testing. Since then, ANNs have been used in the

diagnosis of appendicitis, back pain, dementia, myocardial infarction, acute pulmonary embolism and vasculitis.^{1,3-13} Several urologic applications for ANNs have already been reported. Particular interest of investigators has been focused on several subspecialty fields, including male infertility, uro-oncology, endourology, uro-radiography, and uropathology.⁵

Applications in cytopathology and histopathology

There has been recent interest in the literature in the use of Bayesian belief networks in different areas of pathology including surgical pathology. The main uses of ANNs in histopathology have been as statistical classifiers on data generated by other analytical methods, but there are few studies where digitalized images have been used as the input data for neural networks.^{1,5} Computer programs are commercially available that have been designed to accept variables such as a pathologist's impression of nuclear size and architectural distortion, and to incorporate them in a systematic way to arrive at a diagnosis with a defined degree of confidence (belief). These systems can be adapted to use in different fields of diagnostic pathology.¹³

Histological features for entry can vary depending on a number of factors. For each of the features, a variety of possible grades must first be defined. This can be very flexible; a mixture of 0 to +++, percentage involvement, cell counts, etc, can be used for different variables.¹³ ANNs have been used in the histological diagnosis of different types of breast carcinoma including intraductal forms, parathyroid lesions and hepatocellular carcinoma, histological grading of astrocytomas, prostatic carcinoma, prediction of staging in tumours and automated segmentation of renal biopsies into tubules and interstitium (the ratio of these areas correlate with renal function measured as glomerular filtration rate).

Although, the ANNs have been used most commonly in histopathology, a few studies have also made use of these systems in cytopathology. The two main areas of cytopathology that have been investigated by this system are the screening of cervical smears and breast cytodiagnosis.¹ ANNs have also been used in the cytodiagnosis of pleural and peritoneal effusions, haemopoietic cells, oral epithelial lesions, thyroid lesions, gastric and urothelial lesions.

Applications in renal transplant pathology

The rapid and accurate diagnosis or exclusion of episodes of acute rejection is of vital importance in the proper management of renal transplant patients. Biopsy is the gold standard of diagnosis of acute rejection in most centres. However, if the biopsy is done early in the rejection

process, there can be considerable difficulty in accurate diagnosis of rejection. Banff classification has been developed to standardize the interpretation of renal allograft biopsies.¹⁴ Use of this classification has resulted in improved reproducibility of diagnosis but not significantly, the accuracy. This is because, Banff system uses only two features for diagnosis of rejection; tubulitis and intimal arteritis. However, it has been observed that if more features are included in the evaluation of biopsies, the accuracy of diagnosing rejection is increased. With a large number of variables, all of different significance, it is unrealistic to expect the human brain to correlate the data in a consistent manner. We, Kazi JI et al.¹³ studied BBN (Bayesian belief network) in an attempt to improve the accuracy of diagnosis of early acute renal allograft rejection. This systematic approach increased the sensitivity of detection of early acute rejection (19 out of 21) more than any of the 37 pathologists achieved by conventional histological interpretation (17 out of 21). The study demonstrated the usefulness of inclusion of additional, minor features of acute rejection, such as interstitial edema and eosinophils, in improving the accuracy of diagnosis in these early biopsies beyond the level which is possible using the Banff classification alone (Figure 2).

There are however, certain limitations. Interobserver variation has a significant effect in this system, as demonstrated in our study. Interinstitutional variation will also be a problem, although this was not tested in our system. Most of these problems might be overcome by using a more sophisticated neural network, one which continues to 'learn' and adapts itself to changing characteristics of institutions and the pathologists. Such an evolving network could also

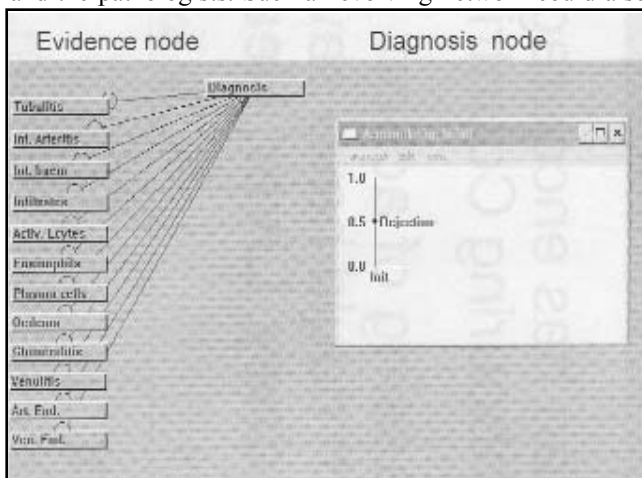


Figure 2. Diagram showing the twelve features used in our study of acute cellular rejection in the evidence mode. Diagnostic mode consisted of either "rejection" or "no rejection".

provide a dynamic measurement of which features were proving to be of most value in the assessment of biopsies. There is also room for inclusion of additional clinical features in the input variables to further enhance the accuracy of diagnosis of early rejection cases in these systems.¹³

Conclusion

It is clear that expert computer systems are going to be increasingly investigated in diagnostic histopathology and cytology in the coming years. Computerized interactive decision support systems, will undoubtedly, lead to improved accuracy, consistency and reproducibility in diagnosis.

References

1. Hamilton PW, Anderson N, Bartels PH, Thompson D. Expert system support using Bayesian belief networks in the diagnosis of fine needle aspiration biopsy specimens of the breast. *J Clin Path* 1994;47:329-36.
2. Montironi R, Whimster WF, Collan Y, Hamilton PW, Thompson D, Bartels PH. How to develop and use a Bayesian Belief Network. *J Clin Path* 1996;49:194-201.
3. Shek TWH. Bayesian Belief Network in histopathology. *J Clin Path* 1996;49:864.
4. Sargent DJ. Comparison of artificial neural networks with other statistical approaches. *Cancer* 2001;91:1636-42.
5. Wei JT, Zhang Z, Barnhill SD, Madyastha KR, Zhang H, Oesterling JE. Understanding artificial neural networks and exploring their potential applications for the practicing urologist. *Urology* 1998;52:161-72.
6. Matsui Y, Utsunomiya N, Ichioka K, Ueda N, Yoshimura K, Arai A, et al. The use of Artificial Neural Network analysis to improve the predictive accuracy of prostate biopsy in the Japanese population *Jpn J Clin Oncol* 2004;34:602-7.
7. Stephan C, Cammann H, Semjonow A, Diamandis EP, Wymenga LF, Lein M, et al. Multicenter Evaluation of an Artificial Neural Network to Increase the Prostate Cancer Detection Rate and Reduce Unnecessary Biopsies. *Clin Chem* 2002; 48:1279-87.
8. Montie JE, Wei JT. Artificial neural networks for prostate carcinoma risk assessment. *Cancer* 2001;91:1647-52.
9. Han M, Snow PB, Epstein JI, Chan TY, Jones KA, Walsh PC, et al. A neural network predicts progression for men with Gleason score 3+4 versus 4+3 tumors after radical prostatectomy. *Urology* 2000;56:994-9.
10. Han M, Snow PB, Brandt JM, Partin AW. Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma. *Cancer* 2001;91:1661-6.
11. Snow PB, Smith DS, Catalona WJ. Artificial neural networks in the diagnosis and prognosis of prostate cancer: a pilot study. *J Urol* 1994;152:1923-6.
12. Babaian RJ, Fritsche H, Ayala A, Bhadkamkar V, Johnston DA, Naccarato W et al. Performance of a neural network in detecting prostate cancer in the prostate-specific antigen reflex range of 2.5 to 4.0 ng/mL. *Urology* 2000;56:1000-6.
13. Kazi JI, Furness PN, Nicholson M. Diagnosis of early acute renal allograft rejection by evaluation of multiple histological features using a Bayesian Belief Network. *J Clin Path* 1998;51:108-13.
14. Racusen LC, Solez K, Colvin RB, Bonsib SM, Castro MC, Cavallo T, et al. The Banff 97 working classification of renal allograft pathology. *Kidney Int* 1999; 55:713-23.