Abstract. The aim of this paper is to provide a snapshot of the application of neural network systems in medical imaging. A few selected case studies are presented, covering the application of neural networks in microscopy imaging in the analysis of cervicovaginal smears and breast cancer histopathology, in ultrasound imaging of the carotid artery, in the analysis of physiological data for lesion placement in pallidotomy and MRI imaging, and in X-ray screening imaging in mammography and chest radiography. It is anticipated that the application of neural network systems in medicine will provide the framework for the development of emerging medical systems, enabling the better delivery of health care.

1. Introduction

The overall objective of Computer Aided Diagnostic (CAD) systems incorporating medical imaging systems or subsystems is to enable the early diagnosis, disease monitoring, and better treatment. The advantages of these systems can be summarized as follows:

- **Standardization.** Diagnoses obtained from different laboratories using similar criteria can be verified.
- **Sensitivity.** Findings on a particular subject may be compared with a database of normal values and/or a decision can be made by a CAD system deciding whether or not an abnormality exists.
- **Specificity.** Findings may be compared with databases for various diseases and/or a decision can be made by the CAD system with respect to the type of abnormality.
- **Equivalence.** Results from a series of examinations of the same patient may be compared to decide whether there is evidence of disease progression or of
response to treatment. In addition, the findings of different CAD systems can be compared to determine which are more sensitive and specific.

- **Efficacy.** The results of different treatments can be more properly evaluated. Medical imaging provides vital information for CAD systems.

The objective of this paper is to present a snapshot of neural network applications in medical imaging systems and how these techniques can be integrated in CAD systems.

According to Haykin [1], a neural network can be defined as follows: “A neural network consist of the interconnection of a large number of nonlinear processing units called neurons; that is, the nonlinearity is distributed throughout the network. We are interested in a particular class of neural networks that learn about their environment in a supervised manner. We have a desired response that provides a target signal, which the neural network tries to approximate during the learning process – achieved by adapting synaptic weights, in a systematic manner.” Furthermore, in the context of signal and image processing applications, neural networks offer the following advantages [1]: nonlinearity, input-output mapping, weak statistical assumptions, learning capability, generalization, fault tolerance, and VLSI implementation.

In this paper, in the next section, the results of literature search on neural networks in medical imaging are presented, followed by a snapshot of selected case studies investigating the application of neural networks in medical imaging systems. These case studies cover the application of neural networks in microscopy imaging in the analysis of cervicovaginal smears and breast cancer histopathology, in ultrasound imaging of the carotid artery, in the analysis of physiological data for lesion placement in pallidotomy and MRI imaging, and in X-ray screening imaging in mammography and chest radiography. Furthermore, in section 4 concluding remarks are presented. This paper shares content with two review papers on adaptive neural networks [2], and data fusion [3] in medical imaging.

2. **Literature review and a snapshot of selected applications**

The INSPEC database was searched with keywords neural networks and medical imaging, and microscopy, or PET, or SPECT, or CT, or ultrasound, or MRI, or X-ray. The number of papers (including both conference and regular journal papers) published under these categories in the years 1991 to 2002 are given in Fig. 1. There were a total of 426 papers published, with 63, 193, and 170 papers published in the

![Fig. 1 Results of the INSPEC database search with keywords neural networks and medical imaging, and microscopy (mic.), or PET, or SPECT (NM for nuclear medicine is given), or CT, or ultrasound (US), or MRI, or X-ray. For each entry, the first, second, and third bars show the hits for the periods 91-95, 96-00, 00-02.](image-url)
years 1991-94, 1995-98, and 1999-02 respectively. These papers cover applications of neural networks for all the physiological systems of the human body, with the majority of them covering the cardiovascular and the nervous systems. Approximately one third of the papers published investigated the application of neural networks in the MRI imaging modality.

A few selected applications of neural networks in medical imaging include: in microscopy imaging in cytogenetics [4], in brain PET for the assessment of epilepsy subjects [5], in CT imaging abdominal organ segmentation [6] and liver tissue characterization [7], in ultrasound imaging in 3D heart motion estimation [8], and diffuse liver disease tissue characterization [9], in MRI imaging [10][11], and in X-ray imaging in hand arthritis assessment [12] and in mammography [13]. Furthermore, more studies investigating the usefulness of neural networks in medical imaging can be found in the following books [14]-[17].

3. Case studies

In this section, case studies of medical imaging applications of neural networks in the following imaging modalities are briefly presented: in microscopy imaging in the analysis of cervicovaginal smears and breast cancer histopathology, in ultrasound imaging of the carotid artery, in the analysis of physiological data for lesion placement in pallidotomy and MRI imaging, and in X-ray screening imaging in mammography and chest radiography.

3.1 The Application of PAPNET in Diagnostic Cytology [18]-[21]

Diagnostic cytology is a branch of pathology that attempts to diagnose human diseases, mainly cancer or precancerous states of various organs, by microscopic examination of cell samples, rather than tissue biopsies. Cells may be obtained by scraping or brushing the surface of the target organs, such as uterine cervix, or by means of a needle syringe for aspiration of fluids accumulated in a body cavity. The differences between benign and malignant cells are reflected via staining procedures of the nucleus of the cells.

Cytologic techniques serve two different purposes: cancer detection and cancer diagnosis. Consider for example the cervicovaginal smear (Papanicolaou test) that has for its purpose the discovery of occult precancerous lesions of the uterine cervix. The screening of cervicovaginal smears is a very difficult human task. Smears may be composed of 50,000 to 250,000 normal cells, with at least 90% of the smear specimens being within normal limits. In a recent survey of American laboratories of cytopathology, the false negative error rate reported ranged from 10 to 20% depending on the abnormality. This measure was recorded for women with biopsy-documented neoplastic lesions.

The PAPNET system has been developed for the purpose of selecting a limited number of cells from cytologic preparations for displaying as images in a high-resolution monitor. The system is interactive and does not attempt to render automated diagnostic opinion. The smear is scanned with a low power objective to identify the areas covered by stained nuclei. The cellular areas are reexamined under the medium power objective that performs the first selection of cells based on size and contrast. The selected regions are reanalyzed with a high power objective and two neural nets to identify images of 64 tiles of single or isolated cells, and 64 images of cell clusters. The ANN assigns a high value to abnormal cells, and a low value to negative cells.
The system was tested on 10 cytopathology labs in the US based on rescreening 497 negative cervicovaginal smears from 228 women who developed biopsy-documented high grade precancerous lesions or invasive carcinoma plus control smears from each laboratory. PAPNET revealed abnormalities that would have led to earlier discovery and treatment of these patients. From the 9666 negative control smears, 127 precancerous lesions were discovered (1.3%). The PAPNET system was approved as a quality control instrument for cervical smears by the FDA.

3.2 A Modular Neural Network System for the Analysis of Nuclei in Histopathological Sections [22]-[26]

The evaluation of immunocytochemically stained histopathological sections presents a complex problem due to many variations that are inherent in the methodology. This subsection describes a modular neural network system that is being used for the detection and classification of breast cancer nuclei named Biopsy Analysis Support System (BASS). The system is based on a modular architecture where the detection and classification stages are independent. Two different methods for the detection of nuclei are being used: the one approach is based on a feed forward neural network (FNN) which uses a block-based singular value decomposition (SVD) of the image, to signal the likelihood of occurrence of nuclei. The other approach consists of a combination of a receptive field filter and a squashing function (RFS), adapting to local image statistics to decide on the presence of nuclei at any particular image location. The classification module of the system is based on a radial basis function neural network. A total of 57 images captured from 41 biopsy slides containing over 8300 nuclei were individually and independently marked by two experts. A five scale grading system, known as diagnostic index, was used to classify the nuclei staining intensities. The experts’ mutual detection sensitivity (SS) and positive predictive value (PPV) were found to be 79% and 77% respectively. The overall joint performance of the FNN and RFS modules were 55% for SS and 82% for PPV. The classification module correctly classified 76% of all nuclei in an independent validation set containing 25 images. In conclusion, this study shows that the BASS system simulates the detection and grading strategies of human experts and it will enable the formulation of more efficient standardization criteria, which will in turn improve the assessment accuracy of histopathological sections.

3.3 A Multi-feature Multi-classifier System for the Classification of Atherosclerotic Carotid Plaques [27]-[29]

There are indications that the morphology of atherosclerotic carotid plaques, obtained by high-resolution ultrasound imaging, has prognostic implications. The objective of this work was to develop a computer-aided system that will facilitate the characterisation of carotid plaques for the identification of individuals with asymptomatic carotid stenosis at risk of stroke. A total of 230 plaque images were collected which were classified into two types: symptomatic because of ipsilateral hemispheric symptoms, or asymptomatic because they were not connected with ipsilateral hemispheric events. Figure 2 shows an ultrasound image of the carotid artery bifurcation with the atherosclerotic plaque highlighted in white colour.

Ten different texture feature sets were extracted from the manually segmented plaque
images using the following algorithms: first order statistics, spatial gray level dependence matrices, gray level difference statistics, neighbourhood gray tone difference matrix, statistical feature matrix, Laws texture energy measures, fractal dimension texture analysis, Fourier power spectrum and shape parameters. For the classification task a modular neural network composed of self-organizing map (SOM) classifiers, and combining techniques based on a confidence measure were used.

The SOM was chosen because it is an unsupervised learning algorithm where the input patterns are freely distributed over the output node matrix. The weights are adapted without supervision in such a way, so that the density distribution of the input data is preserved and represented on the output nodes. This mapping of similar input patterns to output nodes which are close to each other represents a discretisation of the input space, allowing a visualization of the distribution of the input data. The output nodes are usually ordered in a two dimensional grid, and at the end of the training phase, the output nodes are labeled with the class of the majority of the input patterns of the training set, assigned to each node. In the evaluation phase, an input pattern is assigned to the output node with the weight vector closest to the input vector, and it is said to belong to the class label of the winning output node where it has been assigned.

Figure 3 illustrates the distribution of the 160 carotid plaques of the training set on a 10x10 SOM using as input all the 61 texture features.

The SOM classifier yielded also a confidence measure on how reliable the classification result was. The confidence measure was calculated based on the classes of the nearest neighbours on the self organizing map. For this purpose the output nodes in a neighbourhood window centered at the winning node were considered. Several
window sizes were tested. The SOM classifier was first trained with the training set and each output node was labelled with the number of the training input patterns from each class assigned to it. In the evaluation phase a new input pattern was assigned to a winning output node. The number of training input patterns per class assigned to each node in the neighbourhood window around the winning node, were counted. The evaluation input pattern was classified to the class of the majority of the training input patterns. The confidence measure was calculated as the percentage of the majority of the training input patterns to the total number of the training input patterns in the neighbourhood window.

A further enhancement in the calculation of the confidence measure is to give to the output nodes nearest to the winning output node a greater weight than the ones farther away. So a windowing mask is used with the value of the mask depending on the distance of the output node from the winning output node.

In addition, the usefulness of combining neural network classifiers was investigated for the development of medical diagnostic systems. Different feature sets were extracted from the raw data and were inputted into the SOM classifiers. The different classification results were combined using (i) majority voting where the input pattern was assigned to the class with the greater number of votes, and (ii) with the use of the confidence measure where the final classification result was the average of the different confidence measures. The confidence measure was computed as described above and decided the contribution of each feature set to the final result. The idea is that some feature sets may be more successful for specific regions of the input population.

Combining the classification results of the ten SOM classifiers inputted with the ten feature sets improved the classification rate of the individual classifiers, reaching an average diagnostic yield of 73.1%. The same modular system was implemented using the statistical k-nearest neighbour (KNN) classifier. The combined diagnostic yield for the KNN system was 68.8%. The results of this work show that it is possible to identify a group of patients at risk of stroke based on texture features extracted from ultrasound images of carotid plaques. This group of patients can benefit from a carotid endarterectomy whereas other patients will be spared from an unnecessary operation.

3.4 Computer-Aided Diagnosis in Mammography [30]-[34]

Breast cancer is the most common malignancy in women and the second most common cause of death from malignancy in this population. In US, more than 180,000 women develop the disease each year. The practice of mammography is regulated in US through the FDA Mammography Quality Standards Act 92. It is estimated that over 100 groups are working for the detection or characterization of masses and clustered microcalcifications in digital mammography.

Radiographically, mass lesions can be characterised by their degree of spiculation, margin definition, shape, and texture (density, homogeneity). In addition, clustered microcalcifications can be characterized by the morphology of individual calcification, e.g. shape, area, brightness, the heterogeneity of individual features within a cluster, and their spatial distribution. The Department of Radiology at the University of Chicago, developed a CAD system in mammography based on the above mentioned features. Selected features were inputted to an artificial neural network that was trained to compute the percentage of malignancy of suspected lesions. This system, named ImageChecker M1000 was made commercially available by R2 Technology, Inc., Los Alstos, California. A large study involving a retrospective
analysis of 1083 consecutive cancer cases from 13 institutions and more than 24 radiologists was performed as part of the FDA approval process for the ImageChecker M1000. The sensitivity was 98.3% for microcalcification cluster detection and 72% for mass detection with an average false-positive rate of one per image. A prospective component of their study included an analysis of 14,817 cases with the CAD system. No statistically significant change was observed in the radiologists’ workup rate when the system was used as an aid in the screening setting. A second commercial mammography CAD system, SecondLook from Qualia and Scanis, is also seeking FDA approval. Finally, it is noted that in mammographic CAD, the CAD system is used for providing a second opinion and not as a standalone system, and hence need not be perfect.

3.5 Neural Network Analysis of Physiological Data for Lesion Placement in Pallidotomy [35]-[37]

Current pharmacological therapy for Parkinson’s disease loses its usefulness over time. Recently, pallidotomy, a surgical treatment for many of the symptoms of Parkinson’s, has been re-investigated. This procedure requires localization of a small region within the globus pallidus. In this work, a simple electrophysiological analysis, used in conjunction with MRI imaging, provides excellent localization of the target derived from imaging studies alone. Investigation of more complex mathematical analysis may yield additional tools for localization. A new feature in this research is the after-lesioning recordings. It has been proven to be valuable in re-assessing the condition of the patient. If residual activity is observed at the after-lesioning recording “session”, additional lesions might be made, which might further alleviate the Parkinson symptoms.

3.6 A Screening System for the Assessment of Opacity Profusion in Chest Radiographs of Miners with Pneumoconiosis [38]-[40]

The aim of this study was to develop a screening system based on texture analysis for the assessment of opacity profusion in chest radiographs of miners with pneumoconiosis. Chest radiographs were of coal-mine or silica dust exposed miners participating in a health screening program. A total of 236 regions of interest (ROIs) (166, 49, and 21 with profusions of category (shape and size) 0, 1(q), and 1(r), respectively) were identified from 74 digitized chest radiographs by two B-readers. Two different texture feature sets were extracted: spatial gray level dependence matrices (SGLDM), and gray level difference statistics (GLDS). The non-parametric Wilcoxon rank sum test was carried out to compare the different profusion categories versus that of profusion 0. Results showed that significant differences exist (at α=0.05) between 0 vs 1(q), and 0 vs 1(r) for 14, and 12 features respectively. For the screening system, the self-organizing map (SOM), the backpropagation (BP), and the radial basis function (RBF) neural networks classifiers, as well as the statistical k-nearest neighbour (KNN) classifier were used to classify two classes: profusion 0 and profusion 1(q and r). The highest percentage of correct classifications for the evaluation set (116 and 20 cases of profusion 0 and 1(q and r) respectively) was 75% for the BP classifier for the SGLDM feature set. These results compare favorably with inter- and intra-reader variability. In conclusion, texture features provide statistically useful information for the characterization of profusion categories in interstitial lung diseases.
4. Concluding Remarks

Concluding remarks are organized under two main categories: medical imaging issues in general, and adaptive neural networks. For the former, the following remarks are made, motivated by Duncan and Ayche [41]:

- Work in general must be developed and clearly motivated. Analysis should target both normal and pathology cases.
- Medical image analysis tasks are taken in isolation, rather than considered together, i.e. segmentation and registration are pieces of the same underlying task of identifying structure.
- Need to develop appropriate validation and evaluation approaches. There is also, lack of availability of test data sets. Need for the formation of common datasets where algorithms can be compared and contrasted to.
- The medical image analysis community must interact more with other communities, especially the medical physics community.

Some concluding remarks about adaptive neural network applications in medical systems are given for future researchers (see also [42]):

- Clarify the purpose of the study.
- Validate your results appropriately.
- Benchmark against a suitable alternative.
- Develop critical trials involving the exercise of human judgment.

Concluding, neural network applications in medical imaging need to be incorporated into CAD systems, including clinical data, thus enabling the early diagnosis, disease monitoring, better patient treatment, and the offering of a better service to the citizen.

References


