Kestrel Corporation
6624 Gulton Court NE
Albuquerque, NM 87109

Computer-Assisted Chest Radiograph Reader

PI: Peter Soliz, Ph.D.

Grant # 1 R43 OH03595-01

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ART	Adaptive resonance theory
CARRS	Computer Aided Chest Radiograph Reading System
CARRS	Computer-assisted Radiograph Reading System
CT	Computed tomography scans
CWP	Coal workers' pneumoconiosis
DFIDS	Digital Fundus Image Diagnostic System
EpiCC	Epidemiology and Cancer Control Program
GT	Ground truth
HRRC	Human Research Review Committee
ILD	Interstitial lung disease
ILO	International Labor Organization
IRB	Institutional Review Board
LAPART	Laterally-primed adaptive resonance theory
LUT	Look up table
MCMC	Miners' Colfax Medical Center
NEI	National Eye Institute
NIH	National Institute of Health
NIOSH	National Institute of Occupational Safety and Health
RECA	Radiation Exposure Compensation Act
ROI	Region of interest
TIF	Tagged image format
UNM HSC	University of New Mexico Health Sciences Center
VXR	Vidar Film Digitizer

LIST OF FIGURES

LIST OF TABLES

SIGNIFICANT FINDINGS

The specific aim of this Phase I grant was to demonstrate a computer-assisted reader that will reduce inter- and intra-reader variability by employing a state-of-the-art neural network classifier.

Using statistical techniques for characterizing selected regions of interest (ROI) of a chest radiograph, we were successful in classifying tissue types and regions presenting with parenchymal abnormalities. Table 1 and 2 show the results of the tests using a neural network classifier, the laterally primed adaptive resonance theory (LAPART) neural network. The data selected for the study was biased toward the most difficult types of abnormalities to read. The low *kappa* statistics between B-readers and among A- and B-readers demonstrates the issue of inter-reader variability. The *kappa* statistics for the data set selected for the study ranged from 0.17 to 0.51.

Table 1. LAPART classifier results for classifying tissue types and paranchymal abnormal regions in a chest radiograph. GT – ground truth.

GT Total		Normal	Abnormal	Vessel	Rib, etc.
57	Normal	**55**	**2**	**0**	**0**
35	Abnormal	**0**	**31**	**4**	**0**
198	Vessel	**0**	**0**	**198**	**0**
63	Rib, etc.	**0**	**0**	**0**	**63**

Table 2. LAPART classifier results when applied to severity levels in regions with parenchymal abnormalities.

GT Total		1	2	3
96	1	**89**	**7**	**0**
25	2	**2**	**23**	**0**
0	3	**0**	**0**	**0**

Sensitivity and specificity for the first set of tests (Table 1) that were conducted to assess the performance of the neural network and statistical feature characterization program was 0.886 and 0.994, respectively. The tests were aimed at classifying different tissue types (rib, blood vessels, normal tissue) and parenchymal abnormalities. The second set of tests (Table 2) assessed the ability of the statistical features and the neural network to differentiate between severity levels of a region in the chest radiograph. The sensitivity and specificity for the second test was 0.920 and 0.927, respectively.

With strong considerations toward ergonomics, a windows based computer interface was developed and specially designed for use by the physician. No special knowledge of computers or neural networks is required. The system has been designed to be under the control of the user throughout the reading process. The pulmonologist can veto or change the results offered by the computer-neural network. Each ROI selected by the pulmonologist is classified and entered into a database for comparison with other readers and for any future statistical analysis. To make the system affordable, the software was designed to run on an inexpensive Pentium-based personal computer.

Inter- and intra-reader variability would be reduced if one were to accept the results of the computer-aided system without changes. However, knowing that the pulmonologist will want to make some changes, the neural network has a confidence level indicator to show how well an ROI classified by LAPART fits a particular classification pattern. Readers will develop a sense for which ROIs, based on the confidence feedback from the neural network, are more likely to be mis-classified. The high sensitivity and specificity, without human input or adjustments, naturally reduces the inter- and intra-reader variability while improving sensitivity and specificity.

These results indicate that the pulmonologists/readers would benefit greatly from the use of the computer-aided diagnostic system described in this report.

USEFULNESS OF FINDINGS

At this time the findings of the Phase I SBIR project are preliminary and do not apply directly in the prevention of work-place diseases or injuries. The immediate goal of this Phase I project was to demonstrate the potential benefits of a computer-aided system for reading chest radiographs using the ILO protocol.

In the future and after Phase II, when a fully functional and validated computer-aided chest radiograph reading system has been implemented, the impact of such a system will be significant. First, as a screening tool, the computer-aided chest radiograph reading system can be set to a high sensitivity (true positives) and a reasonable specificity (low false alarms). As a screening tool, the computer-aided chest radiograph reading system can reduce physician workload without loss of quality of healthcare. In fact, time not spent on reading normal chest radiographs can be spent on direct contact with the patients. The computer-aided chest radiograph reading system will result in increasing accuracy of the less experienced physician through computer-provided cues and assisted diagonsis of specific regions on a lung.

Perhaps one of the greatest potentials for the computer-aided chest radiograph reading system will be in the added precision of its diagnosis. All parameters calculated from the chest radiographs are repeatable. Quantitative measures of texture, opacity densities, shape, size, *etc.* can be monitored very precisely and longitudinal changes can be detected to much greater sensitivities than is possible by depending on the human vision system. This capability could lead to earlier detection of subtle changes in an individual's lung and could alter the type of care given. For example closer monitoring or limiting exposure of individuals with changes surpassing established thresholds, may prevent progression of disease to more advanced stages.

The chest radiograph is the single most useful tool for clinically evaluating both occupationally related and non-occupational chronic lung diseases, and currently one of the few — and certainly the least expensive and one of the least invasive — means by which lung pathologies can be detected. Chest radiography is the most important diagnostic test available for screening groups of workers for asymptomatic disease. Computer-aided x-ray screening may identify individuals whose disease is minimal. The location and type of opacities seen on a chest radiograph can be a useful tool in diagnosing interstitial lung disease (ILD). The chest radiograph is also an important tool in research on lung diseases and in the medical-legal definition of compensatory disease for worker compensation programs such as the Radiation Exposure Compensation Act program (RECA) and those related to black lung and asbestosis.

The cost of health care is an issue of national concern. An automated or computer-assisted chest radiograph reading system can save the researcher/physician time, eliminate or reduce expensive alternative testing, and aid in reducing inter- or intrareader variability. Quantitatively, the Computer-assisted Radiograph Reading System (CARRS) will improve the diagnostic sensitivity and reliability of radiographs for use in epidemiological research on lung diseases, occupational health screening, and for medical-legal purposes. The proposed system is consistent with the recent emphasis in managed care. The system will be aimed toward accurate detection (none or very low false negatives) for screening, and will offer the radiologist added quantitative information contained in the statistics or texture of the image for improved diagnostic sensitivity/specificity.

The computer-aided chest radiograph reading system to be delivered after Phase II will be validated in an environment which simulates the clinical settings. The goal for the end of Phase I is to have at least one site operating the computer-aided chest radiograph reading system.

ABSTRACT

To reduce inter- and intra-reader variability in diagnosing chest radiographs, a neural network-based computer-aided diagnostic system was developed and tested. The results of an experiment with 124 digitized chest radiographs, demonstrated high degrees of sensitivity and specificity in classifying chest radiographs. The use of a computer-assisted chest radiograph reader eliminated the inconsistencies in the human readers. The Computer-assisted Chest Radiograph Reader System (CARRS) applies recognized principles in the psychophysics of human vision, incorporates neural network-based image analysis and integrates these with a graphical user interface. Advances in digital image processing, and classification techniques have made CARRS feasible for meeting screening, research and development, and clinical requirements.

Through the adoption of the International Labor Organization (ILO) classification procedures, it had been hoped that reader variation in the classification of parenchymal abnormalities could be minimized. The ILO classification of the pneumoconioses is based on a structured procedure for detecting and characterizing patterns on chest radiographs. Numerous studies have shown, however, that inter- and intra-observer variability of radiograph readings by trained medical personnel has persisted [1].

The methodology was implemented through the following tasks: 1) From a data base of several thousand patients, a set of 205 chest radiographs were manually graded by two pulmonologists; 124 of the films were then digitized at 12-bit high spatial resolution. 2) Textural features were calculated using high order statistical techniques [2]. The features were classified by the pulmonologists to "train" a neural network to extract classification rules chest radiographs based on the ILO methodology. 3) The neural network classification from the graded system was tested using 65 chest radiographs.

For 5-10 areas selected by the pulmonologist on the chest radiograph, a feature vector composed of image characteristics such as density distribution, entropy, fractal dimension, opacity counts, shape, *etc.* was calculated. This feature vector characterized numerically the areas used by the pulmonologist to grade the radiograph. The neural network trained on the same regions used by the pulmonologist, and through a quantitative feature vector, "learned" the characteristics of each ILO classification.

The laterally-primed adaptive resonance theory (LAPART) neural network was selected [3, 4]. LAPART presents the results of the training in human interpretable "rules." LAPART trains in a single pass ("fast"), making it attractive for a clinical setting. To demonstrate that calculated characteristics of chest radiographs could be used to train the LAPART to classify the radiographs using objective and quantitative parameters, a pilot study was conducted. 32 statistical features from 124 digitized chest radiographs were calculated. Statistical parameters were selected based on their contribution to the separation of classes. The parameters included entropy, contrast, fractal dimension, and co-occurrence statistics. The computational efficiency of LAPART was demonstrated in this experiment. LAPART was trained on a Pentium 200. Training time for 10 experiments, where the training and testing radiographs were randomly selected, averaged 15 seconds per experiment.

In conclusion these experiments showed that LAPART is a worthy candidate for basing a semi-automatic chest radiograph classification system. The experiments demonstrated that CARRS has low classification variability and has a significantly high accuracy. The results of each of the 10 experiments showed that the LAPART neural network algorithm could be trained to "learn" the extent of the "hypercubes" which represented each of the six classifications in 32-dimensional space. LAPART classification accuracy averaged 98 %. "Truth" was determined by the two radiologists.

A. Background.

Respiratory diseases have a tremendous public health impact in terms of loss of quality of life and mortality. An equally large factor is the economic impact that results from the loss of work days and the financial drain on the health care system from the often intensive and long-term care required by these patients. Lung cancer, a leading cause of cancer deaths in the U.S. [5], is also the most common cause of cancer deaths worldwide, a result of the high incidence rate and low survival rate. An estimated 87% of lung cancer cases are attributable to cigarette smoking [6]. Occupational exposures are estimated to contribute to the additional 13% [7].

Chronic lung diseases, such as the pneumoconioses, one of which is silicosis, and interstitial lung disease (ILD) or pulmonary interstitial fibrosis, were the fourth leading cause of death in 1995, responsible for approximately 4% of all deaths in the United States [8]. The primary consequence of chronic lung disease is dyspnea, or pathologic breathlessness [9]. The results are lifestyle restrictions, impaired respiratory tract clearance mechanisms, excessive mucus production, and reduced lung capacity that all contribute to more frequent, severe, and prolonged respiratory infections, loss of work-time, and decreased quality of life. These diseases also require lengthy hospitalizations. Coal workers' pneumoconiosis (CWP) is identified by a specific pattern of x-ray abnormalities and a history of exposure to coal dust. Between 1979 and 1990, almost 30,000 deaths were attributed to CWP. Chronic lung diseases have both occupational and non-occupational causes [10]. The most common causes of occupational ILD are inhaled free silica, asbestos, mixed dust, coal, beryllium, and cobalt. Non-occupational ILD can result from many conditions, including collagen vascular diseases such as rheumatoid arthritis, systemic lupus erythematosus, and mixed connective tissue disease; from drugs and treatments such as antibiotics, cardiac and chemotherapeutic drugs, and radiation; as a result of viral, bacterial, or fungal infections, pulmonary granulomatosis, and idiopathic pulmonary fibrosis [11]. The ILDs are a heterogeneous group of disorders that comprise more than 130 entities [12]. The prevalence of all ILD in the U.S. has been estimated to be approximately 20-40 cases per 100,000 members of the population, and has been reported in 15% of the patients seen by pulmonary physicians [9]. A recent population-based study suggested that the prevalence of the disease may be higher and that these diseases may often go unrecognized [13].

The chest radiograph is the single most useful tool for clinically evaluating both occupationally related and non-occupational chronic lung diseases, and currently one of the few — and certainly the least expensive and one of the least invasive — means by which lung cancer can be detected [14]. Chest radiography is the most important diagnostic test available for screening groups of workers for asymptomatic disease. X-ray screening may identify individuals whose disease is minimal [5]. By limiting additional exposure, it may be possible to prevent progression of disease to more advanced stages. Asbestos, another major form of occupationally related lung disease, resulting in approximately 10,000 deaths annually, presents with characteristic x-ray abnormalities. In addition, the location and type of opacities seen on a chest radiograph can be a useful tool in diagnosing ILD. The chest radiograph is also an important tool in research on lung diseases and in the medical-legal definition of compensatory disease for worker compensation programs such as the Radiation Exposure Compensation Act program (RECA) and those related to black lung and asbestosis [12]. However, its role in screening or in the epidemiological study of lung diseases is limited, primarily by technical considerations [15].

Radiography is a primary component of respiratory disease surveillance, and standardization is an important aspect of surveillance radiography [16]. The 1980 International Labor Organization (ILO) system was developed in an attempt to standardize the interpretation of pneumoconioses. It is a scheme that was designed specifically to systematically record radiographic changes that reflect the inhalation of ust [14, 17]. The original intent was to facilitate epidemiologic comparisons of radiologic abnormalities, not to define pathologic entities [18]. More recently, however, use of the ILO Classification System has been extended to include clinical diagnosis of all ILDs, as well as litigation in cases of occupationally and environmentally induced lung disease [15, 16, 19-22]. The ILO system, although a major step toward

quantifying the disease as observed in chest radiographs, is necessarily based on less than totally quantitative measures used to describe profusion and to estimate sizes of features like opacities. The profusion of opacities is scored based on the observations of a trained B-reader. The reader assesses the profusion of opacities compared with ILO standard radiographs. The classification recognizes the existence of a continuum of the progression of the features, from no opacity to advanced stages [23]. From the twelve possible categories, the readers are asked to give a discrete score ranging from 0/-, 0/0, 0/1, 1/0, 1/1, *etc.* to 2/3, 3/3, and finally 3/+. The profusion of opacities is scored based on a mentally integrated value from three areas of each lung: upper, middle, and lower. Findings at the low end of the scale (0/1, 1/1) are often equivocal, inconsistent, and nonspecific. For this reason, such results are often not useful for diagnosing disease in individuals [13].

Reader Variability Issues.

Through the adoption of the ILO classification procedures, it had been hoped that reader variation in the classification of parenchymal abnormalities could be minimized. Because the ILO classification of the pneumoconioses is based on a structured procedure for detecting and characterizing patterns on chest radiographs, this system has the potential for limiting the variations in interpretation. Numerous studies have shown, however, that inter- and intraobserver variability of radiograph readings by trained medical personnel has persisted [1, 22]. Mulloy *et al.* examined 134 investigations of pneumoconiosis in four peer-reviewed journals from 1985 to 1990 and identified a number of potential problems with the application of the ILO classification system: the need for multiple readers to quantify inter- and intrareader variability; the lack of repeatable analysis; the documentation of formal training of the readers; the use of different versions of the ILO system; the blinding of readers to workers' exposures, and others. Though more sensitive in resolving smaller opacities, chest CT scans do not entirely solve the problem of reader variability, and standards for interpretation, similar to the ILO system, are just being developed.

Reader variability has been reported as being a significant problem affecting classification consistency of pneumoconiotic radiographs, *e.g,.* Amandus *et al.* [24] and Ducatman *et al.* [22]. Interreader variability is well documented, as is the variability in the same reader on different occasions [25]. Muir *et al.* [26] studied an alternative approach which required the reader to view and assess six lung zones separately in classifying chest radiographs. This approach relieved the reader of having to integrate mentally the reading for the entire lung into a single value. Muir *et al.* [26] showed statistically that overall agreement between readers was better when they were asked to read zonally, and that disagreement was greater when readers classified the chest radiographs in a single reading, forcing the mental integration process.

The Computer-assisted Reader eliminates many of these problems or quantify the reader's interpretation process by documenting pathological features in terms of measurable parameters such as density, density distributions, area coverage, size and shape of opacities, location, number of opacities, *etc.* The Computer-assisted Reader will implement the time-tested ILO process as a computer-assisted system which applies the same criteria as the present ILO system, but uses quantitative measures to give the reader more precise information regarding the pathological features being analyzed. The basis for a reader's classification can be readily studied by other readers or researchers who can use the classified radiographs in quantitative, statistically-based studies. Longitudinal studies will benefit from having precise location, profusion, and density measurements for comparison of the progression or resolution of the observed and measured pathological features. Statistics on individual readers or classes and features can be collected and used to perform mathematical assessments of reader variability.

Phase II.

An end-to-end system, like the Computer-assisted Chest Radiograph Reader system, which not only provides a means for entering diagnostic data, and viewing and storing chest radiographs, but objectively and thoroughly quantifies the statistical nature of the chest radiographs and associated lung pathologies (interstitial diseases, silicosis, lung nodules, *etc.*) will be prototyped and fully tested with a large population data base before the conclusion of Phase II. The proposed system will use off-the-shelf

computer interface technology to minimize development and product costs and will be integrated into an existing PC-based medical image-based diagnostic environment and researcher/physician user interface which is currently being developed under other project funding. Computer technology based on Intel's Pentium™ processor presents an opportunity for using inexpensive, off-the-shelf, state-of-the-art computer hardware technology. Our image processing tools are, likewise, state-of-the-art algorithms developed for similar medical image classification applications.

A comprehensive study will be performed in Phase II using the Miners' Colfax Medical Center chest radiograph database and B-readers from the University of New Mexico Health Sciences Center, Epidemiology and Cancer Control (EpiCC) Program. The study will demonstrate the sensitivity and specificity of the Computer-assisted Chest Radiograph Grading System. With several thousand chest radiographs already having been graded by two or more B-graders, the Computer-assisted Grading System will be applied to these data to demonstrate its effectiveness in reducing intra- and inter-reader variability, for producing greater quantitative data, and for adding to the statistical significance of many epidemiological studies.

Commercial.

The cost of health care is an issue of national concern. An automated or computer-assisted chest radiograph reading system can save the researcher/physician time, eliminate or reduce expensive alternative testing, and aid in reducing inter- or intrareader variability. Quantitatively, the Computer-assisted Radiograph Reading System (CARRS) will improve the diagnostic sensitivity and reliability of radiographs for use in epidemiological research on lung diseases, occupational health screening, and for medical-legal purposes. The proposed system is consistent with the recent emphasis in managed care. The system will be aimed toward accurate detection (none or very low false negatives) for screening, and will offer the radiologist added quantitative information contained in the statistics or texture of the image for improved diagnostic sensitivity/specificity.

Clearly there is commercial potential in radiology for state-of-the-art chest radiograph classification technology being developed by the pattern recognition, image processing and neural network communities. Phase I explored the feasibility of using advanced image processing and neural networks to improve chest radiograph classification and diagnostic capabilities. In Phase II, a program with the University of New Mexico Health Sciences Center will be implemented to perform a statistically based study to assess the efficacy of the CARRS in reducing reader variability and for quantifying features associated with occupationally related pneumoconioses.

The future commercial system would be an end-to-end system, incorporating an automated or computer-assisted screening system, a system for clinical and field use, and a research tool for integrating quantitative information into the fundamental understanding of pathologies currently detected and documented by radiologists. The Computer-assisted Reader prototype system developed in Phase I and used for the pilot study of Phase I will require only minor modifications for the larger population study that will be conducted in Phase II. A commercialization plan will address technical, clinical, and economic issues, such as the device's potential for reducing healthcare costs by offering a superior means for screening chest radiographs and telemedicine communications before suggesting more expensive procedures like computed tomography (CT) scans. We are working with a number of venture capitalists and manufacturers who are aware of our work. These commercialization and manufacturing components will be added to the team between Phase I and Phase II and will support the proposal preparation for Phase II. Commercialization is the ultimate goal of this project.

Related Research.

In 1974 Kruger, *et al.* [27] recognized the potential of computer-aided reading of chest radiographs for reducing future reading loads on manual systems. They suggested mass screenings using computer techniques for implementing a version of the ILO protocol. Kruger, *et al.* observed that most textures can

be described by their size, shape, orientation, and repetition rate. Their experiments with statistical techniques to define the characteristics of a texture in the lung form the basis of a number of other studies that followed including our own. Computer technology of the time limited their experiments to coarse spatial resolutions (9.8 pixels per mm) and ROI sizes under 100 x 100. In spite of the limitations imposed by the technology available to Kruger, the results of the computer-based diagnosis was comparable, though slightly less accurate than the physician's diagnosis.

In another paper, Ledley, *et al.* [28] demonstrated a texture analysis technique for automatic analysis and classification of pneumoconioses (black lung disease). They provide an excellent review of the concept of texture in the context of spectral analysis. As with the other early research into computer-aided diagnosis and texture analysis of chest radiographs, these researchers were handicapped by the lack of easily available, high powered computers, such as are often found in a clinic or even a home or office.

Monnier-Cholley, *et al.* [29] evaluated the improvements to diagnostic accuracy in detection of interstitial opacities on chest radiographs by employing a computer-aided system. Their results show a statistically significant improvement for residents and attending radiologists. For example the area under the ROC curve improved from 0.943 to 0.969 for residents and from 0.960 to 0.972 for the radiologists. The feature vector for analyzing the texture and pattern in the radiographs was based on Fourier techniques. Using the power spectrum of the lung texture patterns, two features are calculated – namely the root mean square variation and the first moment of the power spectrum. These two features represent the magnitude and coarseness of the infiltrate. Symbols that were related to the features were superimposed on the digitized radiographs and presented to the radiologists in the computer-aided portion of the tests.

Haralick, *et al.* [30] give a detailed description of the statistical techniques adopted for this study for texture analysis. Katsuragawa, *et al.* [31] describe an automated technique for determining the physical characteristics of lung textures in chest radiographs. Their purpose was to develop a computer-aided diagnostic system for interstitial lung diseases. They chose to use the power spectrum technique for analyzing the texture. Like other studies, the two features extracted from the power spectrum were root mean square variations and the first moment of the power spectrum. Katsuragawa shows only the feature extraction results and a qualitative presentation of the classification of normal and abnormal based on the two features. A rigorous statistical analysis is not presented to show the ROC or sensitivity and specificity of the classifier based on the two features. Nevertheless, qualitatively, the results do demonstrate a potential for classifying normal and abnormal regions of a lung using this technique. Kim and Park [32] applied statistical texture analysis techniques to characterize microcalcifications in mammograms. The feature vectors were classified using a back-propagation neural network.

Neural networks have been used by a number of researachers to classify radiographs. Penedo, *et al.* [33] developed implemented a back-propagation neural network as part of computer-aided diagnostic architecture for detecting lung cancer nodules in chest radiographs. A feature vector that quantifies the *curvature peak space* was used in the neural network classifier. *Curvature peak space* is the local curvature of the intensity surface.

Lo, *et al.* [34] developed and tested a neural network that first "learns" to recognize patterns in the lung tissue through a network that simulates some aspects of the human vision system. This part of the system is referred to as a neocognitron neural network [35]. The output of the neocognitron produces the feature vector which is processed by a back-propogation neural network. This technique has also been applied by the principal investigator for the project. It is a very slow learning neural network and was dismissed for use in feature detection in the Phase I.

Computer-based techniques have been developed in mammography that can be applied to the automated analysis of chest radiographs. Schnorrenberg, *et al.* [36] analyzed breast cancer biopsies using receptive field filters to enhance the cell images. Savol, *et al.* [37] took the approach of growing nodules in a given region. This technique has been adopted for our research as one of the algorithms for segmenting an image under the control of the user. It was not applied extensively because of its strong need for manual

intervention. Nevertheless, the technique may be useful for special purposes when the patterns in the ROI show clear tendencies toward high and low intensity patterns. We plan to explore its use in Phase II. Giger, *et al.* [38] used differencing approaches to accentuate lung nodules. Their approach processed a single projection in two different ways. One process maximizes the signal-to-noise of the nodule while a second image processing filter minimizes the nodules. For both processes the background remains essentially unchanged, so that the difference of the two resulting images produces a difference map that highlights the nodules. This technique will be incorporated into the next version of our own feature extraction methodology. Giger's work showed a good detection rate with low false alarms. No sensitivity or specificity values were given.

B. Phase I Specific Aims.

The specific objective of this Phase I grant was to demonstrate a computer-assisted reader that will reduce inter- and intra-reader variability by employing an existing, state-of-the-art neural network classification process. A prototype system, Computer-assisted Chest Radiograph Reader System (CARRS), based on Kestrel's Computer-assisted Medical Image Diagnostic System [23, 39-41], was implemented on a low-cost Pentium-based computer, in a graphically oriented Windows® environment. Standard point and click menus are used to interface a specially designed chest radiograph image manipulation interface. The system applies recognized principles in the psychophysics of human vision, incorporates neural network-based image analysis and integrates these through rapid prototyping of graphical user interfaces for low-cost personal computers. Advances in computer-based image processing, and data fusion techniques have made computer-assisted radiograph classification feasible for meeting screening, epidemiological research, and clinical requirements. Additionally, quantitative classification and analysis of radiograph images can yield information critical for characterizing the nature of the present state and longitudinal progression of lung diseases such as pneumoconioses.

Digital image analysis has advanced dramatically over the past decade and most recently with the introduction of high-throughput personal computers such as those based on the Intel Pentium™ processor. Advanced image processing capability is available at very affordable prices. This is especially true when compared to costs for storage, archiving, and management of film-based radiographic images. Digital image processing and analysis has the added advantages of having greater sensitivity to image density (intensity) variations than human vision, and has the ability to perform precise calculations of feature characteristics that can provide a precise record of the size, profusion, and texture of large and small opacities. Electronic transmission of digital images has become routine and is particularly advantageous because specialists are often centrally located and otherwise not easily accessible to clinics or remote sites. Spatial resolution matching, which is attainable by traditional film-based processes, is now possible.

The specific aim of this Phase I research project was to demonstrate a significant reduction of inter- and intra-reader variability through the integration of a computer-assisted system for implementing ILO chest radiograph grading. Specifically, the aim was met by performing the following tasks.

- A set of 200 chest radiographs from the Miners' Colfax Medical Center Outreach Program data base of several thousand patients, was selected and digitized at 12-bit high spatial resolution. The radiographs had previously been graded by two NIOSH-certified B-readers, Dr. David Coultas and Dr. David James.
- Textural features that were calculated using high order statistical techniques were combined with features identified by the two B-readers to "train" a neural network developed by Dr. Peter Soliz, Principal Investigator, to extract classification rules for reading chest radiographs based on the ILO methodology.
- The CARRS neural network classification was tested using independent sets of chest radiographs. The measure of success was a quantitatively derived set of statistics showing the degree of reduction in inter- and intra-reader variability.

The approach for meeting each of these specific objectives is described in Section C. The innovations of the CARRS are summarized below:

- Integration of current practice, i.e. B-reader analysis and interpretation based on standardized ILO methodology, with computer-based quantification techniques.
- Integration into the Computer-assisted Chest Radiograph Reader System state-of-the-art image processing, pattern recognition, and neural network-based classification algorithms with modern Windows™ interfaces, voice recognition commands, and low-cost personal computer technology.
- Demonstration of a neural network-based algorithm as the basis for reducing inter- and intra-reader variability through a computer-assisted chest radiograph classification system that emulates current ILO practices.

C. Procedures and Methodology.

The data. Since 1989, the Miners' Colfax Medical Center (MCMC) Outreach Program, in collaboration with pulmonary physicians from the UNM Health Sciences Center, has offered free screening for mining-related diseases to active and retired miners in the southwestern United States. Persons who have worked in mines of any type are eligible for this program. Participants are examined at the MCMC in Raton, NM or in a mobile clinic that periodically visits mining communities in New Mexico and south central Colorado. Archived posterior-anterior chest radiographs from this program were used to investigate the nature and presentation of pneumoconioses in approximately 200 miners. Radiographs from the cohort were analyzed by two NIOSH-certified B-readers and served as a benchmark for assessing the feasibility of a partially or fully automated radiograph image analysis system. One of the consultants to this project, Dr. David James, serves as the Medical Director for the MCMC Outreach Program and is a faculty member at the University of New Mexico Health Sciences Center, Institutional Review Board (IRB). Therefore, review of the protocol for use of the radiograph set and protection of patient confidentiality was done by the University of New Mexico Health Sciences Center Human Research Review Committee (HRRC). The overall goal of this project was to achieve significant gains in sensitivity and specificity in classifying and screening chest radiographs, and to demonstrate the appropriateness of computer-based image processing technology for eventual commercialization in world-wide use.

As noted earlier, reader variability is a significant problem in the interpretation of radiographs for pneumoconiosis. Many revisions of the ILO classification system focused on reader variation. One revision in particular, the introduction of a set of standard radiographs, made it easier to require conformity with a national standard of reading by interpreters of radiographs. ILO standard radiographs of small opacities depict the center of each major category of profusion. When interpreting a radiograph, the reader is asked to make a judgement about which of the standard films a particular radiograph most closely matches. For the Phase I study, the ILO standard radiograph set was scanned (see below for scanning procedure) to provide a "gold standard" for training the CARRS.

Several databases of chest x-rays were available to perform this study. The Miners' Colfax Medical Center Outreach Clinic database consists of thousands of chest radiographs. The database was sufficient to find the necessary distribution of the disease progression to test the prototype system. UNM Health Sciences Center has established a population-based registry of patients with ILD in Bernalillo County, New Mexico. Chest x-rays for these patients are all read by two NIOSH_certified B readers. A database of approximately 10,000 chest radiographs taken as part of the pre-employment and ongoing health surveillance physicals for Grants Mineral Belt uranium miners is available for future use in the development and test of the Computer-assisted Reader.

Data Scanning.

Data Scanning of all chest radiographs followed a protocol developed to maximize the amount of information acquired during the digitization process. To maximize the entropy of each image (information content) each image was prescanned with a Vidar VXR Film Digitizer to calculate an improved look-up table (LUT) and then scanned three times to yield a final digitized image. All images

were scanned at 300 dpi (dots per inch) at a depth of 12 bits. For a 14 inch square radiograph scanned at this setting, the pixel size was 85 microns.

Rationale for scan protocol: Once an image is digitized the amount of information (grayscale, contrast, texture) inherent in the digital file becomes fixed. Adjusting a histogram after scanning only enables the viewer to better perceive the recorded information. By creating a system for scanning, we were able to change the way the VXR worked, manipulating the analog to digital signal ratio based on previous prescans and optimizing the stored visual information before the data was fixed by the final scan. Three TIF format images were created in this process: a twelve bit image using the entire film as the basis for adjusting the scan, a twelve bit image using only a region of interest (ROI) as the basis for adjusting the scan, and an eight bit image for record keeping.

Software provided with the Vidar VXR Film Digitizer contained translation tables. The tables were used to convert or map data from the film digitizer into a range better suited to the Phase I viewing requirements. The goal was to give the readers (physicians) a digital image of a chest radiograph, which optimized, or enhanced what can be seen on an actual film when viewed on a light table. Two translation tables were used to generate the 12 bit images: a Histogram Equalized translation table, and a Linear Histogram Independent table. The Histogram Equalized translation table evens out the number of pixels in the light and dark areas of the film based on the histogram of the last prescan. It also adds contrast to the image. The Linear Histogram Independent translation table maps the specified low grayscale value to grayscale zero and the specified high grayscale value to the highest grayscale value (value 4095 in 12 bit), and with Dark Enhance disabled, results in the darkest possible image of all the default translation tables available with the Vidar VXR. When used, both of these translation tables will differ from one film to another; therefore, a prescan must be made first to produce a translation table for each film. The steps for using the tables performed in the following sequence on each film:

1. Set the translation table to Histgm eqlzed.
2. Set Save to disk File format to TIF, do not name the file to be saved
3. Press Prescan
4. Next, select an ROI for only right lung (to avoid heart)
5. Record x,y coordinates and width, height information
6. Set the translation table to Linear HI with the upper limit of 4095 (12 bit depth)
7. Press Prescan
8. Turn off Auto size
9. Enter the ROI parameters just recorded (Top, or Y1, Left, or X1, Width and Height)
10. Press Scan (note: at this point the operator is making an actual scan of the entire film using a translation table made from the lung ROI only for that image)
11. Set the translation table to Histgm eqlzed
12. Create table, save as 12base file name for that image.tif
13. Press Save
14. Turn Auto size ON
15. Press Scan (this action applies the newly recorded translation table to the scan)
16. Set the translation table to Linear HI, with the upper limit of 4095 (12 bit depth)
17. Prescan
18. Set the translation table to Histgm eqlzed
19. Create table, save as 12base name for that image.tif
20. Press Save
21. Press Scan (note: at this point the operator is making an actual scan of the entire film for that image using a translation table made of the entire film)
22. Set the translation table to Linear HI with an upper limit of 255 (8 bit depth)
23. Press prescan
24. Set the translation table to Histgm eqlzd
25. Create table, save as 8base name for that image.tif
26. Press Save
27. Eject film

At the end of this process the user had scanned and saved three TIF format files. Two 12 bit files and one 8 bit file. The twelve-bit file, which used the translation table created from the lung ROI had the best percentage and distribution of grayscale values, and was the optimal image.

Basis for Comparison – Manual Reading.

B-Readers.

Table 3 and Table 4 show the distribution of the profusion levels assigned by the two principal B-readers, Drs. Coultas and James, to radiographs archived in the MCMC Outreach Program database

Table 3. Radiographs with Irregular Opacities, Miners' Colfax Medical Center Outreach Clinic.

Reader #1 ↓	Reader #2 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**1574**	113	35	12	2	-
0/1	21	**6**	7	3	-	-
1/0	8	3	**9**	11	-	-
1/1 - 1/2	6	-	1	**7**	7	-
2/1 - 2/3	1	-	-	-	**4**	1
3/2 - 3/+	1	-	-	-	-	**4**

The level of inter-reader variability that exists even between these two experienced B-readers is evident in Table 3. Of the total radiographs that were classified for the profusion level of irregular opacities, the readers agreed on 87 %. However, for the 153 radiographs that were classified by either as 0/1, there was agreement on only 6, or 4 %. Similarly, for the 74 radiographs that were classified by one or the other reader as a 1/0, they agreed on only 9, or about 12 %. The kappa statistic, κ, was described by J.L. Fleiss **[42]** as a means for measuring agreement among readers. κ statistic for Table 3 is 0.34.

Table 4. Radiographs with Rounded Opacities, Miners' Colfax Medical Center Outreach Clinic.

Reader #1 ↓	Reader #2 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**1616**	77	13	2	-	1
0/1	31	**24**	18	-	-	-
1/0	14	7	**27**	12	3	-
1/1 - 1/2	6	1	11	**29**	17	-
2/1 - 2/3	-	-	-	6	**19**	2
3/2 - 3/+	-	-	-	-	-	-

A similar level of inter-reader variability is apparent in the classification of profusion levels in radiographs with rounded opacities. κ statistic for Table 4 is 0.51.

A set of 210 radiographs was selected from the Miners' Colfax Medical Center Outreach Program database and digitized. The radiographs that were selected did not reflect the distribution of the larger population, which was heavily biased toward the 0/0 profusion category. The Phase I data set contained a greater percentage of "non-normal" radiographs (*i.e.*, profusion category of 0/1 or higher) than was contained in the overall population. The data set was purposefully biased toward radiographs where there was disagreement between the B-reader interpretations. In addition, a larger number of 0/1 and 1/0 films were selected than is reflected in the original population distribution. The latter profusion categories are over-represented on the B reader certification examination because of the high level of variability between readers in categorizing films into these categories. It was our hypothesis that CARRS could be most helpful in classifying films that fell into these "borderline" profusion categories. Only radiographs of good or acceptable film quality were selected.

Table 5 illustrates the level of inter-reader variability of the B-readers for the 210 radiograph set. Of the total radiographs that were classified, the readers agreed on 38.6%. However, for the 103 radiographs that were classified by either as 0/1, there was agreement on only 16, or 15.5 %. Similarly, for the 52 radiographs that were classified by one or the other reader as a 1/0, they agreed on only 12, or about 23.1%. κ statistic for the readers in this table is 0.21.

Table 5. Radiographs used in study with Rounded Opacities, Miners' Colfax Medical Center Outreach Clinic.

Reader #1 ↓	Reader #2 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	25	16	3	2	0	0
0/1	57	16	7	0	0	0
1/0	6	6	12	8	0	0
1/1-1/2	1	1	7	14	1	0
2/1-2/3	1	0	3	8	15	0
3/2-3/+	0	0	0	0	2	1

Class:	% agmnt	Total	agrmnt
0/0	**22.5%**	111	25
0/1	**15.5%**	103	16
1/0	**23.1%**	52	12
1/1-1/2	**30.0%**	40	12
2/1-2/3	**50.0%**	30	15
3/2-3/+	**33.3%**	3	1
total:	**38.6%**	210	81
w/o 0/0	**34.1%**		

A-Readers.

92 of the radiographs were also graded by the two A-readers, a matrix comparing the grades assigned by A-readers were produced. For example, Table 4 shows the results of the two A-certified readers' (Drs. L. Ketai and K. Mulloy, UNM HSC Cancer Research & Treatment Center) classification for rounded opacities for a set of 106 radiographs.

Table 6 illustrates the level of interreader variability for the 92 radiographs. Of the total radiographs that were classified, the readers agreed on 50 %. However, for the 34 radiographs that were classified by either as 0/1, there was agreement on only 4, or 11.8 %. Similarly, for the 26 radiographs that were classified by one or the other reader as a 1/0, they agreed on only 4, or about 15.4%. κ statistic is 0.30.

Table 6. A-reader results from 92 radiographs used in study with Rounded Opacities, Miners' Colfax Medical Center Outreach Clinic.

Reader 3 ↓	Reader 4 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	31	4	0	0	0	0
0/1	10	4	6	0	0	0
1/0	5	7	4	2	0	0
1/1-1/2	1	3	2	3	2	0
2/1-2/3	0	0	0	2	4	2
3/2-3/+	0	0	0	0	0	0

Class:	% agmnt	Total	agrmnt
0/0	**60.8%**	51	31
0/1	**11.8%**	34	4
1/0	**15.4%**	26	4
1/1-1/2	**20.0%**	15	3
2/1-2/3	**40.0%**	10	4
3/2-3/+	**0.0%**	2	0
total:	**50.0%**	92	46
w/o 0/0	**26.3%**		κ=.30

Table 7 illustrates the level of inter-reader variability for the 96 radiographs graded by both A-readers. Of the total radiographs that were classified, the B-readers agreed on 43 %. However, for the 45 radiographs that were classified by either as 0/1, there was agreement on only 7, or 15.6 %. Similarly, for the 26 radiographs that were classified by one or the other reader as a 1/0, they agreed on only 8, or about 31 %. κ statistic is 0.24.

Table 7 96 x-rays read by B-readers that are in common with the A-readers.

Table 7. B-reader results from radiographs in common to both A-readers.

Reader 1 ↓	Reader 2 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**17**	6	0	2	0	0
0/1	24	**7**	3	0	0	0
1/0	4	4	**8**	3	0	0
1/1-1/2	0	1	3	**6**	0	0
2/1-2/3	0	0	1	2	**5**	0
3/2-3/+	0	0	0	0	0	**0**

Class:	% agmnt	Total	agrmnt
0/0	**32.1%**	53	17
0/1	**15.6%**	45	7
1/0	**30.8%**	26	8
1/1-1/2	**35.3%**	17	6
2/1-2/3	**62.5%**	8	5
3/2-3/+		0	0
total:	**44.8%**	96	43
w/o 0/0	**36.6%**		

In general, the experience of those evaluating the ILO classification system has been that readers with less experience tend to produce more "anomalous" results when compared to "expert" readers [18] and also tend to reader higher levels of profusion than their more experienced counterparts [43]. Thus, one recommendation for reducing inter-reader variability has been to require that a minimum number of films be read per year as a condition of maintaining B reader status [18].

Table 8 through Table 10 illustrate the level of A- versus B-inter-reader variability for the 106 radiographs. Of the total radiographs that were classified, the A- and B-readers agreed on about 50%.

Table 8. B-reader 1 versus A-reader 4 results from radiographs in common.

Reader 1 ↓	Reader 4 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**35**	9	*1*	0	0	0
0/1	13	**2**	*3*	0	0	0
1/0	2	8	***3***	2	0	0
1/1-1/2	0	2	*4*	**3**	4	0
2/1-2/3	0	0	*0*	2	**2**	1
3/2-3/+	0	0	*0*	0	0	**1**

Class:	% agmnt	Total	Agrmnt
0/0	**58.3%**	60	35
0/1	**5.4%**	37	2
1/0	**13.0%**	23	3
1/1-1/2	**17.6%**	17	3
2/1-2/3	**22.2%**	9	2
3/2-3/+	**50.0%**	2	1
total:	**47.4%**		46
w/o 0/0	**21.2%**		κ =.24

Table 9. B-reader 2 versus A-reader 4 results from radiographs in common.

Reader 2 ↓	Reader 4 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**21**	2	*1*	0	1	0
0/1	24	**8**	*2*	0	0	0
1/0	5	9	***3***	2	0	0
1/1-1/2	0	2	*5*	**2**	1	0
2/1-2/3	0	0	*0*	3	**4**	1
3/2-3/+	0	0	*0*	0	0	**0**

Class:	% agmnt	Total	agrmnt
0/0	**38.9%**	54	21
0/1	**17.0%**	47	8
1/0	**11.1%**	27	3
1/1-1/2	**13.3%**	15	2
2/1-2/3	**40.0%**	10	4
3/2-3/+	**0.0%**	1	0
total:	**39.6%**		38
w/o 0/0	**23.9%**		κ =.17

Table 10. B-reader 1 versus A-reader 3 results from radiographs in common.

Reader 1 ↓	Reader 3 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**29**	3	0	0	0	0
0/1	5	**8**	4	2	0	0
1/0	5	3	**5**	2	0	0
1/1-1/2	1	0	2	**5**	0	0
2/1-2/3	0	0	0	2	**5**	1
3/2-3/+	0	0	0	0	0	**0**
	40	14	11	11	5	1

Class:	% agmnt	Total	agrmnt
0/0	**67.4%**	43	29
0/1	**32.0%**	25	8
1/0	**23.8%**	21	5
1/1-1/2	**35.7%**	14	5
2/1-2/3	**62.5%**	8	5
3/2-3/+	**0.0%**	1	0
total:	**63.4%**		52
w/o 0/0	**46.0%**		κ =.49

Table 11. B-reader 2 versus A-reader 3 results from radiographs in common.

Reader 2 ↓	Reader 3 → 0/0	0/1	1/0	1/1 - 1/2	2/1 - 2/3	3/2 - 3/+
0/0	**16**	16	0	0	0	0
0/1	2	**8**	7	2	0	0
1/0	1	3	**5**	6	0	0
1/1-1/2	2	0	4	**0**	1	0
2/1-2/3	0	0	0	0	**7**	0
3/2-3/+	0	0	0	0	0	**0**

% agree on:		74	70
0/0	**43.2%**	37	16
0/1	**21.1%**	38	8
1/0	**19.2%**	26	5
1/1-1/2	**0.0%**	15	0
2/1-2/3	**87.5%**	8	7
3/2-3/+		0	0
total:	**45.0%**		36
w/o 0/0	**41.7%**		κ =.27

Feature Selection.

Although our aim was to show that computer-assisted techniques can reduce significantly inter- and intra-reader variability, our approach and associated techniques did not dispense with the requirement for highly trained radiologists/pulmonologists to provide the cues related to location of features. We referred to this activity as region of interest (ROI) selection. Once the ROI was identified by the pulmonologist, its textural characteristics were quantified through the application of image characterization algorithms.

Using the digitized images and film radiographs, the B-readers, selected ROIs from each of the chest radiographs, then evaluated and entered into the database a type and severity for each ROI. The two B-readers performed the task together in order to have full agreement and establish "ground truth."

Parenchymal abnormalities consistent with pneumoconioses were graded according to the ILO standard methods. A computer interface was developed to allow the pulmonologists to circumscribe on a digitized image the area on the radiograph presenting small or large opacities. The precise location, area involved, density, and a statistical representation of the texture, shape and size was calculated and recorded.

The chest radiograph depicts lung tissue as well as other nonuniform anatomical structures such as the rib cage and internal organs. The pathological features of interest are superimposed on this background noise. Regions of interest (ROI) were selected manually to collect the data on specific structures, such as small opacities, large opacities, inter-rib spaces, ribs, *etc.* The abnormality, if one was present, was detected and classified by the B-reader. Since the two B-readers performed this task together, using both digital images and the original plain film, the data were considered as the basis for all other future comparisons. This baseline labeling and classification will be refereed to as "ground truth." To develop and test any computer-based classifier, it is necessary to create a data set comprised of examples of each class.

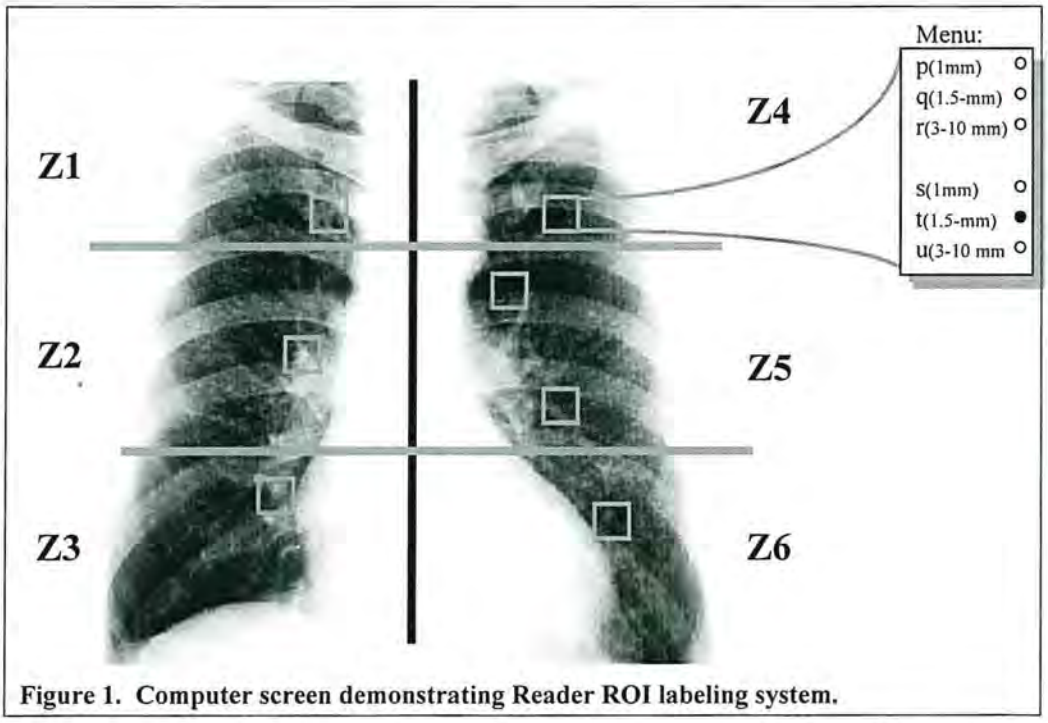

Figure 1. Computer screen demonstrating Reader ROI labeling system.

To provide examplars of the different types (p, q, ...) and severities (1, 2, 3) of rounded opacities, the two B-readers were asked to identify and classify regions in the 124 x-rays which presented with these pathologies. A computer program, an element; of CARRS, was developed to assist the pulmonologist with reading the x-rays, selecting each ROI, and recording the ROI size, position, and diagnosis. Figure 1 presents an example of an x-ray which has been "read" by the two B-readers. The examplars are a requirement for all classifiers including the neural network that was employed in our research. An element of the CARRS was developed that allowed the pulmonologist to "point-and-click" on a digitized chest radiograph that was present on a computer monitor. The pulmonologists were asked to identify the affected region, define its margin and classify it by type and severity. Examples of normal tissue were also collected.

A total of 963 ROIs were collected and labeled. Table 12 shows the distribution of examplars used in the Training and Testing of the neural network classifier.

Table 12. Summary of ROIs selected and "read" by the two B-readers.

TOTAL	Normal	Parench.	Vessel	Ribs
963	170	83	556	154
100%	18%	9%	58%	16%
			Total ROIs	963
Type ↓ & Severity→ (Total by type)		**1**	**2**	**3**
p	(3)	3	0	0
q	(69)	56	12	1
r	(2)	0	1	1
s	(0)	0	0	0
t	(8)	3	2	3
u	(1)	1	0	0

The ROIs selected were characterized using technique based on sum and difference histogram, or co-occurrence matrix statistics [30, 44] parameters, which represent the texture, contrast, density, *etc.*

ROI Parameters.

Once the pulomonologist has identified an ROI and the disease progression has been classified according to the ILO methodology, there is a need to quantify features in the ROI and relate the features to the classification of a radiograph. For purposes of computer-based classification the ROI must be characterized in some quantitative and repeatable manner. To detect and characterize various lung pathologies, we have developed and tested a number of feature characterization algorithms. The numerous patterns and complex variations of the different areas of the lung/chest were characterized quantitatively using features derived from a co-occurrence matrix.

Texture analysis played a very important role in the analysis of the chest radiographs. Texture is a property of the spatial distribution of gray levels in a local area of the image, and will have some extent greater than the local property of a pixel. Texture attributes are described qualitatively in terms of coarseness, roughness, and directionality. Quantitatively, texture may be characterized through statistical measures, such as contrast, entropy, and homogeneity.

A second approach for texture analysis, the Laws [45] texture energy method applies a convolution to the image with a set of small masks. The local texture energy results from the convolution of the mask with the image. Fractals offer one quantitative means for representing the natural shape, size, and intensity of features present in a radiograph. Three-dimensional fractal models have been successfully applied to the problems of texture segmentation and classification, estimation of 3-D shape information, and distinguishing between perceptually "smooth" and perceptually "textured" or "rough" surfaces in an image [46-49]. With fractals, one is able to describe a surface or shape with a dimension larger than its topological dimension. In other words, the fractal dimension of a surface corresponds to an intuitive notion of roughness. Fractal dimensions of the radiographs were computed using a program based on a method by Keller *et al.* [50]. Clearly, as the perceived roughness increases, so does the fractal dimension, demonstrating that fractals can represent our (the radiologist's) notion of roughness, perhaps relating to the profusion of small opacities.

A statistically-based approach, derived from the co-occurance matrix [2], was the principal technique used to quantify the texture for use by the CARRS. The gray level co-occurance matrix statistics were defined at several pixel displacement statistics. By characterizing the texture within an ROI at several pixel separations, it is hoped that a textural feature with a multiple pixel harmonic or cycle will be

captured by on e of the displacement distances. One interpretation of texture is that gives by a set of statistical measures of the spatial variations of gray backing an image. Although there are a number of approaches for determining the spatial statistics, including auto correlation [51], power spectrum [52], Markov random field [53], fractals [54], and others, we chose to use the co-occurance matrix technique [55].

Texture information in an image may be revealed through an analysis of the statistical relationships that may exist in the spatial distribution of gray levels [30]. A set of co-occurance matrices $P(i, j)_{d,0}$, can be defined which characterize the nature of the texture in a gray-level image. The indices (i, j) represent frequencies that gray level i and grey level j are found a distance, d, apart in the direction given by 0.

With definitions of features based on the co-occurance matrix, 7 features were calculated for 4 pixel separation distances. The texture features were assumed to have no directional dependence. Although 4096 gray levels are captured in the scanning process, not all features exploited the detailed of the 12-bit gray level representation. Table 13 shows the features calculated, the pixel separation (d) and the effective gray levels used for each feature set.

Table 13. Feature vector was composed selecting different combinations of pixel separations, d, and gray level resolution.

	Pixel Separation				Grey Levels			
Feature	1	2	4	8	64	128	256	4096
Contrast	X		X				X	X
Correlation	X	X	X	X	X	X	X	
Entropy	X		X		X	X	X	X
Homogeneity	X	X	X	X			X	
Shade	X		X	X			X	
Prominence	X		X	X			X	

Contrast is given by

$$CNTRS(d) = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (i-j)^2 \, P(i,j)_d$$

Where N_G is the number of gray levels. Contrast is the difference of gray levels squared, weighted by the co-occurrence (probability) matrix. This means that when the contrast value is low, the concentration of co-occurrences is high along the main diagonal indicating a coarse texture.

Correlation is given by

$$CORR(d) = \frac{\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} (ij) \, P(i,j)_d - \mu_x \, \mu_y}{\sigma_x \, \sigma_y}$$

Where μ_x and μ_y are mean values and σ_x, σ_y are standard deviations. Correlation represents the linear dependencies of gray levels.

Entropy is given by

$$ENTROPY_d = -\sum_{i=1}^{N_G} \sum_{j=1}^{N_G} P(i,j)_d \log P(i,j)_d$$

Entropy is a measure of the disorder in the scene. Entropy achieves a maximum when all P(i, j)s are equal.

Local homogeneity is given by

$$HOMO\,G(d) = -\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}\frac{P(i,j)_d}{1+(i,j)^2}$$

This feature represents the measure of the amount of local similarity in the scene.

Shade difference is given by

$$SHADE\_d\,(d) = \frac{\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}\left[(i-j)-(\mu_x-\mu_y\right]^3 P(i,j)_d}{\left(\sigma_x^2 - 2\sigma_{x,y} + \sigma_y^2\right)^{\frac{3}{2}}}$$

Prominence is given by

$$PROM\,(d) = \frac{\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}\left[(i-j)-(\mu_x-\mu_y\right]^4 P(i,j)_d}{\left(\sigma_x^2 - 2\sigma_{x,y} + \sigma_y^2\right)^2}$$

The N-dimensional space represented by these features was explored to determine the set with the greatest sensitivity and specificity for classifying chest radiographs. A non-parametric technique based on neural networks was selected.

Neural Network Description.

Artificial neural networks have been found to perform tasks in diagnosis of some medical images with results that match or exceed those of human radiologists [56, 57]. Neural networks have several significant advantages over humans -- speed, consistency, higher dimensional memory, and longer temporal memory. Neural networks, however, have been limited by the need to "train" using complex and unstable learning rules and extensive databases. Training is usually a computer-intensive process requiring extensive analysis by the engineer and numerous retraining cycles to assure "convergence" to a near optimal solution. Training the neural networks has been the domain of neural network experts or physicians with significant in-depth knowledge of neural networks. The decision process of the trained neural network is not normally amenable to inspection or interpretation by the radiologist, who must accept the "output decision" with little or no insight into the rationale for the outcome. These neural networks fall into the class of "multilayer perceptrons."

The center piece of the CARRS is the laterally primed adaptive resonance theory (LAPART) neural network. In cybernetics, machine vision, computational intelligence, bionics, and numerous other fields, biological metaphors have been used to solve technical problems in the respective areas of study. Similarly, neural network models have been proposed which in some limited ways emulate the human or biological vision systems in "learning" to recognize from shape, texture, and spectra. To perform the categorization of the individual pixels into classes or areas which share common spectral characteristics, a neural network which is being used extensively by the PI, Dr. Soliz, in pattern recognition problems was applied to the chest radiographs.

The research team successfully developed a neural network which avoids the deficiencies of multilayer perceptrons in the classification of pathological features in retinal images. While nearly all the major neural networks were considered, our choice was a self-organizing network based on the adaptive resonance theory (ART) developed by Carpenter and Grossberg [4]. The advantages and attributes of ART that we exploited are discussed below. A variant of the ART neural network (LAPART for laterally primed adaptive resonance theory) by Caudell [58] has been routinely applied to developing "rules" from

large data stores such as medical history databases (records and imagery). The LAPART neural network trains in a single presentation of the data. Medical images are analyzed wherein LAPART is trained on tens of millions of data points from images in a matter of 2-3 minutes on a Pentium PC. Phase I demonstrated these and other attributes which make the LAPART an important new tool in medicine in general and for diagnosing chest radiographs in particular. The ART is a self-organizing artificial neural network that uses features such as the density, texture, shape, *etc.* to categorize a region according to the combinations of values for those parameters. A short summary of the LAPART neural network algorithm and the rationale for the selection of the LAPART neural network is given. Details on the theory behind the ART ANN may be found in the literature, including Grossberg [59] and Healy and Caudell [60]. The LAPART network:

- Adapts quickly and efficiently to the variability of different radiograph image systems.
- Does not degrade to unacceptable levels under some conditions such as missing parameters (channels, noisy pixels, *etc.*), *i.e.*, no singularities.
- Produces results in "real-time", *i.e.*, is not computationally intensive or algorithmically complex.
- Is sufficiently flexible to integrate and to embed into a number of different types of computer workstations or other systems.
- Is partitioned easily to take advantage of current and next generation parallel processors.

The LAPART has a much greater ability than most other techniques to capture patterns in data. Without the need for complex, explicit algorithms, it can represent patterns in spatial and temporal data. The LAPART has several computational properties that are not generally found as a complete set in other neural networks:

- The training process is less time consuming than that for back propagation. Retraining is not necessary when new data are presented to a "trained" LAPART.
- A vigilance parameter allows one to set the coarseness of the classification. This means one can train the LAPART to discover arbitrarily finer or coarser classes based on any set of measured feature parameters. New categories are easily integrated into a trained network.
- Learning is assured to be stable and normally occurs in one iteration [3]. Long and often unsuccessful learning trials, as with the multilayer perceptron neural network, can reduce the appeal and even impact the clinical implementation of a neural network.

The learning concept for the ART architecture is described in references by Carpenter. Briefly, the sets of calculated and measured parameters comprise a data vector systematically presented to the LAPART neural network. The LAPART stores combinations of feature patterns. Each input pattern, representing a region of the radiograph, is presented to LAPART. The most highly activated recognition node (the "most similar" node) in LAPART is selected as the matching class. If the "best" class is not a sufficiently close match to the input pattern, a different recognition node (the next most active) then becomes the choice node. If no match exists, a new class or template is formed. This means that a new input pixel pattern not sufficiently fitting any existing pattern has been presented to the LAPART neural network. The LAPART responds by creating a feature category that maps to the defined set of feature categories. This LAPART model has been used extensively by Dr. Soliz and UNM for the past four years, is well tested, and is known to be suited for the task of developing a pathological feature classification system based on input parameters.

In the application of the LAPART to classifying chest radiographs, one could think of the templates as representing sets of ranges in N-parameter space. For example, if template 1 represented opacity parameter, template 2, texture, and template 3 entropy, then in three-space the clusters could represent a specific ILO classification. By adjusting the vigilance parameter, ρ, one could change the resolution, *i.e.*, number of clusters so that each hypercube represented an ILO class.

Figure 2 illustrates how LAPART's output templates may be represented. Although the feature space would be N-dimensional (N=number of features used in the classification), illustrated in Figure 2 is for N=3, where the dimensions of the feature vector are entropy, density and fractal dimension. The LAPART will be trained by using the features calculated from each of the 12 standard ILO images. Each cube (template) would contain the pixels with similar "features". These cubes will be given a class, such as the ILO class, 0/1, 1/0, etc., or an intermediate classification which is meaningful to the pulmonologist, such as low, medium, high profusion, or as used in the ILO standard form: q, r, s, etc. If desired, the templates may be presented with a false color display, much like other modalities (PET, *etc.*). LAPART "sorts out" the widely varying calculated and human selected parameters into "hypercubes" of self-consistent feature regimes. Figure 2 (left) illustrates the output of Radiograph Classes from the LAPART, within the LAPART. LAPART discovers classes of radiograph regimes based on the combined values of the coded input vector.

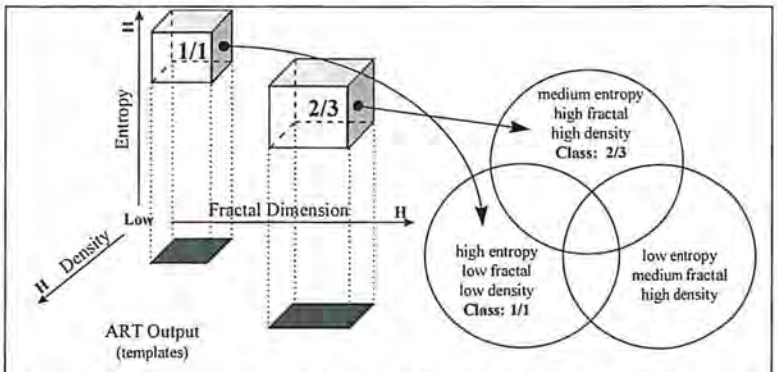

Figure 2. Concept used by ART for training on a feature vector (depicted by three features: entropy, density, fractal dimension). Each classification will be represented by the LAPART as a N-dimensional cube where extent of the cube is determined by the range of values for a given feature.

LAPART systems perform pattern recognition by comparing an input pattern with template (or object class prototype) patterns stored in the network's synaptic memory. Each template W_J is a pattern of synaptic strengths that can be compared parameter by parameter with an input pattern $\boldsymbol{I}$. In LAPART, a dual set of feedforward and feedback connections to/from an input layer of network nodes to a layer of recognition nodes stores template W_J.

The input pattern activates recognition nodes through the feedforward subset of connections for all **J**, and the most highly activated recognition node (the "choice" node) reads out its stored template W_J through its subset of feedback connections to the input layer. If the feedback pattern is a sufficient match to the input pattern, most of the activated input nodes, representing the binary 1's in $\boldsymbol{I}$, remain active; if not, the mismatch between input and feedback quenches the activity of most input nodes, allowing the vigilance node of the network to reset the recognition layer. This results in a different recognition node (the next most active) becoming the choice node. A sufficient match of W_J with $\boldsymbol{I}$ (resonance) for some **J** results in a stable episode of activation of the current choice node. During this episode, the choice node's connection pattern to/from the input layer changes to adapt to I, thereby modifying the template. If no such **J** exists, a recognition node whose connections have not yet formed a template has its connections to the input layer modified to form a new template $W_J = \boldsymbol{I}$.

In the notation of Carpenter *et al.* [4], the modified binary template pattern is the intersection of the subset of 1's in $\boldsymbol{I}$ with the subset of 1's in the resonating template W_J. If the pattern components are ordered and represented as vectors, we denote the update at resonance by

$$W_J^{new} = \mathrm{I} \cap W_J^{old} \qquad (1)$$

where a component of $\boldsymbol{I} \cap W_J^{old}$ is 1 if and only if the corresponding binary components of $\boldsymbol{I}$ and W_J^{old} are both 1, and 0 otherwise. The choice node is selected from among eligible nodes in the recognition layer (those that have not undergone reset with the current input $\boldsymbol{I}$) by a competitive interaction that effectively selects J to maximize the quantity

$$\frac{|I \cap W_J|}{\alpha + |W_J|} \qquad (2)$$

Here, the $|\bullet|$ operation produces the number of 1's in a binary pattern $\boldsymbol{I}$ and α is a small, positive constant for the ART network. Resonance occurs if $\boldsymbol{I}$ and the choice template W_J are sufficiently matched, measured by the test

$$\frac{|I \cap W_J|}{|I|} \geq \rho \qquad (3)$$

where ρ is the vigilance of the network, $0 \leq \rho \leq 1$.

The LAPART system with m input nodes (m binary components per input pattern), setting

$$\alpha \leq \frac{1}{m - 1} \qquad (4)$$

ensures that a maximal subset template W_J relative to an input pattern $\boldsymbol{I}$, if one exists in the ART system, is the strongest attractor for $\boldsymbol{I}$, in the sense of (2). A maximal subset template relative to $\boldsymbol{I}$ is the largest template such that

$$I \cap W_J = W_J \qquad (5)$$

W_J^{new} is a maximal subset template relative to $\boldsymbol{I}$ immediately following the template update in (1). Learning in an LAPART system self-stabilizes for a given set of input patterns when each pattern in the set is immediately recognized by the system and no template recoding takes place. That is, (3) is satisfied by the first choice selected according to (2) and $W_J^{new} = W_J^{old}$ in (1). LAPART thus has the potential for rapid, stable, adaptive pattern recognition.

A large data set consisting of 300 hundred or more chest radiographs will be used in the training. The task of training on hundreds or several thousand images is not out of the question for the LAPART neural network. It has been observed that as theory predicts, LAPART always converged in two training epochs and ran extremely rapidly on a Pentium equipped personal computer.

Interface Design, Implementation, and Operation.

The computer-aided radiograph reading systems (CARRS) was developed using the same principles, technology, and some of the same analytical software previously developed for NIH. A digital fundus image diagnostic system (DFIDS) was developed by Dr. Soliz, under NEI Grant EY10951-03, to be a general biomedical image processing system for facilitating the study of digital biomedical images and the programming and testing of new biomedical image analysis algorithms[61]. One element of DFIDS, LAPART, was applied directly to classification of the chest radiograph images. The motivation for developing CARRS and DFIDS has been to reduce the present high cost of commercial image processing systems by exploiting the rapid growth in software tools for the development of graphical user interfaces and accompanying means for rapid prototyping of commercial systems. Working with UNM pulmonologists, Dr. Soliz used the CARRS to apply state-of-the-art image analysis and pattern recognition techniques to chest radiographs images.

The approach used in CARRS integrates the human analyst into the classification. It exploits those human attributes which allow people to readily detect and recognize patterns while leaving those aspects of the problem which require precise measurement and counting to the computer. In this respect CARRS takes the approach of presenting the data to the computer so as to invoke the computer's quantitative abilities, while allowing the human expert to guide the process and assure the accuracy of the results. CARRS combines a neural network with interactive analysis to provide a tool to assist the pulmonologists in assessing and quantifying the parchymal pathologies. The system provides information not available in

the current reading protocols, including the precise location of each lesion, texture, density, and size and shape parameters. Each lesion is registered, or mapped into a global coordinate system, so that an individual lesion can be re-called and studied by the same or other readers. The system can compute detailed statistics for each patient, and to assist lateral studies in which a patient and even a particular lesion are followed through time.

The system uses a multi-tiered approach. The work flow is depicted in Figure 3 below. The user selects the chest radiograph that will be read using CARRS, and from that, a number of local Regions of Interest (ROIs) to present to the neural network classifier. The ROIs will represent regions between the ribs that are thought to present opacities. The LAPART neural network clusters a set of numerical values that characterize each ROI according to there textural, density, *etc.* clusters. The "trained" LAPART associates the combination of texture, density, *etc.* to a previously presented ROI with similar features. The previously presented ROI is from the training data set and will have a classification label, such as type (q, r, s, *etc.*) and severity (1, 2, or 3). The user then enters into a dialogue with the system, using a graphical user interface to assign interpretations to the CARRS suggested ROI classification according to a menu of options.

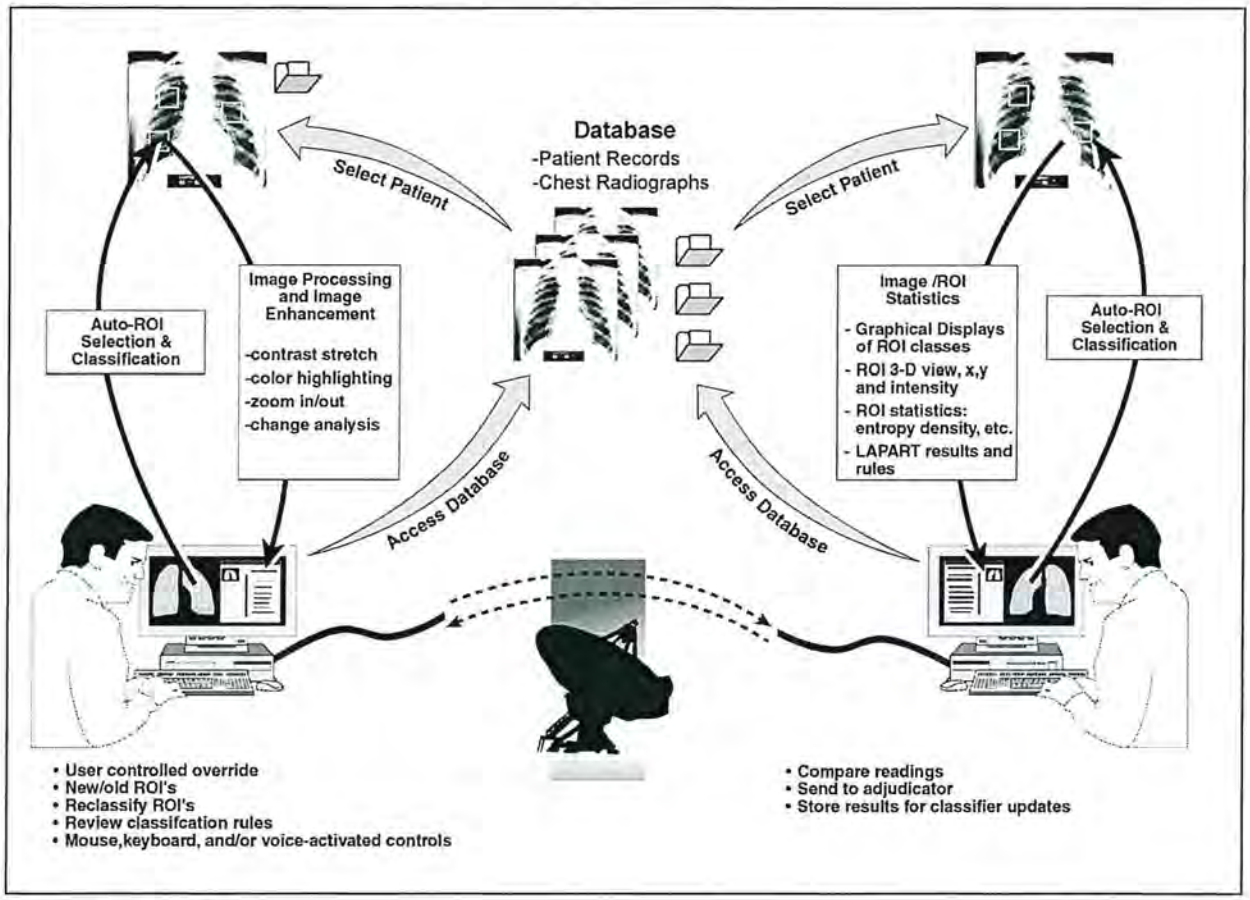

Figure 3. CARRS flow.

The classification is being performed by LAPART. Through the generation of a human – friendly rule set, the LAPART provides total insight to the pulmonologist's into its classification process. For example, one of the CARRS graphical displays shows each feature (density, etc.) used in the classification, its range of values for each type and severity, and the specific value for the ROI being analyzed. Each of these steps is described in greater detail below.

Preprocessing ROIs with LAPART.

It is well-known in pattern recognition that the performance of a classification or clustering routine is also affected by the prior probabilities: Bayesian theory tells us that a ROI is much more likely to be classified correctly if the probability of the correct class was high. CARRS allows the user to select a ROI to exclude extraneous features and improve the prior probabilities. The user selects the shape, size and location of the ROI.

The ability of a classification routine to separate an image into meaningful clusters is seriously impaired when there is inter- and intra-subject variations in the image. Chest radiograph imaging invariably introduces these variations, due to the difficulties of positioning and instrument. The first step in normalizing the images as much as possible was in the care and diligence applied to the scanning and digitization of the images. The techniques were described above.

The sets of calculated and measured radiograph image-derived parameters compose a data vector systematically presented to the ART neural network. The ART stores combinations of textural, density, *etc.* features input templates (patterns). ART systems perform pattern recognition by comparing an input pattern with template (or object class prototype) patterns stored in the network's synaptic memory. Each template is a pattern of synaptic strengths that can be compared parameter by parameter with an input pattern. In ART, a set of feedforward and feedback connections to/from an input layer of network nodes to a layer of recognition nodes stores the template.

In the results section below we depict how the ART operates on the data and how the classification of combinations of pixel parameters is presented by ART. The large dimensionality of the input data vector makes it difficult to illustrate outputs on two-dimensional charts. The large dimensionality does not necessarily impair ART's ability to "sort out" the widely varying pixel parameters into "hypercubes" of self-consistent ROI-characteristic regimes. ART discovers classes of ROI characteristics based on the combined values of the input vector. In a supervised version two ARTs, training simultaneously, produce a mapping which associates ROI-characteristic regimes with reader-specified features [3].

We chose to work with the LAPART rather than with other, more popular techniques such as the back propagation neural network because of its efficiency and rapid convergence characteristic which made the LAPART more suitable for performing large numbers of numerical experiments. The LAPART neural network simulates some aspects of human learning and memory which are exploited in the application. For example, unlike the backpropagation, which must be retrained if it is to incorporate new data (which might include a never-before-encountered chest radiograph), the LAPART class of neural network which we use has been shown to converge rapidly and simply creates a new category and the accompanying weights in order to add a new class. LAPART was designed to treat "out-liers" (novelty detection), such as those which occur when a piece of data with a significant error (poor quality x-ray) is entered, or when an unusual natural feature appears [3, 60]. The backpropagation and other techniques do not deal with this situation gracefully.

The ART was implemented as a means for clustering and classifying complex ROI feature vectors. LAPART simultaneously maps ROI feature values into two clusters: one is the vector composed of a type (including normal tissue types and opacity types); the other is a vector of severity. LAPART forms an association between the predictors (input feature vector) and the output vector used to represent the type and severity.

Figure 4 illustrates the learning paradigm for the LAPART architecture. In a manner to be detailed above, the left processing element, ART_A, of the LAPART neural network stores combinations of input templates (features such as contrast, entropy, etc.). The right processing element, ART_B, stores predict and vectors which hold the resulting patterns, for example, the ILO-based classification of type and severity. The LAPART learns to associate clusters of input vectors with clusters of output vectors.

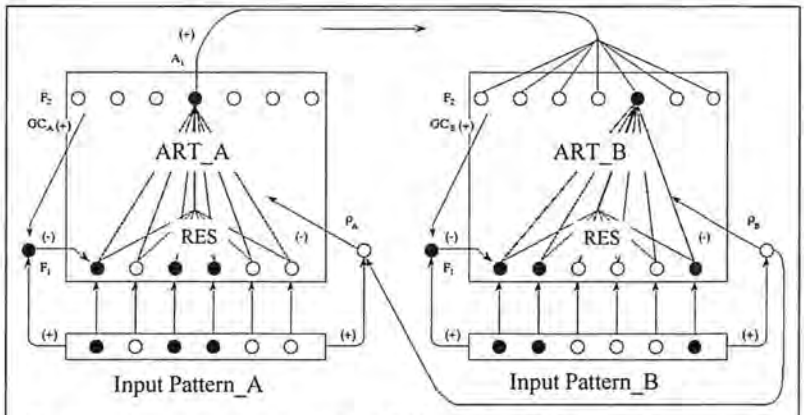

Figure 4. A supervised ART, LAPART, Architecture as implemented by Caudell [58].

Adaptive resonance theory systems perform pattern recognition by comparing an input pattern with template (or object class prototype) patterns stored in the network's synaptic memory. Each template is a pattern of synaptic strengths that can be compared parameter by parameter with an input pattern. In ART, a dual set of feedforward and feedback connections to/from an input layer of network nodes to a layer of recognition nodes stores the template.

The Data Representation.

The nature of the problem is central to formulating the structure of the neural network and the feature vector. Figure 5 depicts the approach for presenting the ROI feature vectors to the LAPART. In applying the LAPART to the chest radiograph images to classify type and severity, we presented the ART_A module with a feature vector, *e.g.* from a given ROI, x_0 to day x_n. ART_A would "learn" pattern of the features, x_0 to x_n and provide output clusters of similar ROI feature patterns. The ROI feature patterns are the parameters which are used to characterize quantitatively the nature of the ROI. Simultaneously, ART_B would classify "consequences", *i.e.* predicted type and severity which result from the input sequence to ART_A. By adjusting the vigilance parameter, ρ, one could change the resolution, or number of clusters, so that ART_A or ART_B classes could be divided into as many or as few classes as the application requires.

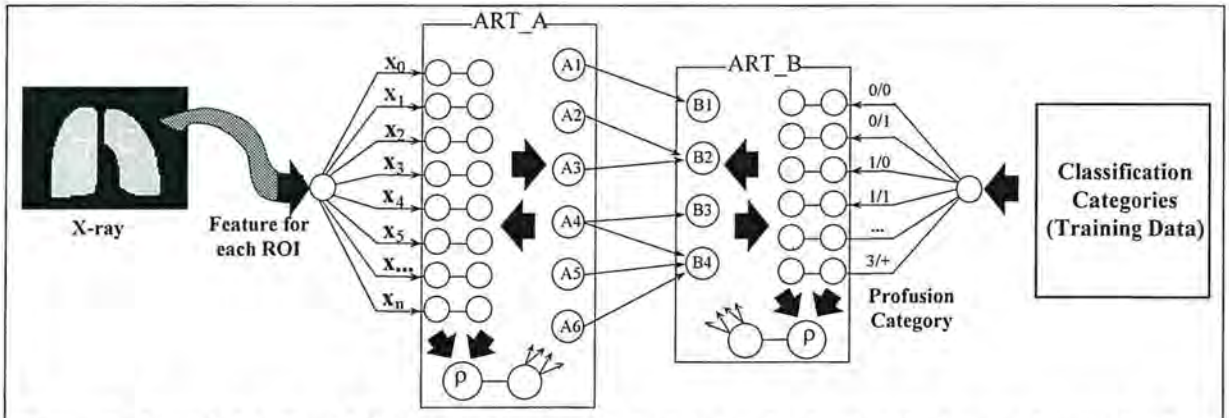

Figure 5. Depicted is the LAPART architecture for training and inferring the type and severity of an abnormality from previously "learned" patterns.

Figure 5 illustrates the LAPART architecture for training on a set of ROI feature vector data. Typically, one uses a large training data set that would have representative values for each type and severity. On the left is the ROI feature vector. Tens to hundreds of features, each of which is a different mathematical representation of the nature of the ROI's texture, density, opacities (size and shape), are presented to the ART_A module. The ART_A classifies the feature vector $\{x_0, x_1, x_2, \ldots x_n\}$, *e.g.* {entropy, density, . . . } with other "similar" feature patterns. Similarity is defined by the value of vigilance,ρ, that is used during the training. For $\rho = 1$, patterns in a class would be identical. For settings of $\rho < 1$, the patterns will match in proportion to the value of ρ. For example, at $\rho = 0.5$, the patterns in a class will have 50 % of their sequences in common. At the same time that the ART_A is producing a map that correlates classes or labels that are presented to ART_B for classification as the predicted pattern.

A mapping occurs between the two ART modules. A pattern in ART_A implies a classification or category in ART_B. The analysis of these mappings will determine the complexity and uniqueness of the feature dectors and these classes. One-to-one (*e.g.*, A1 → B1) or many-to-one (*e.g.*, A2 & A3 → B2) mappings between ART_A classes and ART_B classes signifies stable classifiers. One-to-many (*e.g.*, A4 → B3 & B4) or many-to-many would be ambiguous, unpredictable states. LAPART does not produce one-to-many mappings.

Experiment Design and Controls.

Figure 3 presented the flow chart for the data processing through the Computer-aided Chest Radiograph Reading experiment. ROIs were selected for all 200+ chest radiographs. The two principal B-readers "read" the images using the film and the digital formats. This process provided the ROIs and the classification of type and severity for each ROI. Co-occurance matrix-derived features were calculated for each ROI. Using these features and the B-reader provided diagnosis, we selected 90% of the chest radiographs to train the neural network. The remaining 10% of the chest radiographs were used to perform the test of CARRS.

In the test of CARRS the "rules" generated by the trained LAPART neural network, were applied in "real-time" to classify a chest radiograph. To start the test, the first of the 10% of the digitized chest radiographs that were withheld from the training is presented to one of the B-readersThae computer screed presents the chest x-ray along with the LAPART prediction to the pulmonologist. An ROI is highlighted and in a few milliseconds the LAPART has classified the ROI into one of 3 normal categories or 18 categories of type (6) and severity (3). The CARRS presents the output window (right side of Figure 1). The output is the classification of the ROI. The pulmonologist may accept the LAPART classification or change it. The pulmonologist number him/her decision then by clicking onOK goes to the next ROI. When the pulmonologist is satisfied that the most relevant areas have been diagnosed, the ROI classes are integrated into a single, overall chest radiograph diagnosis.

D. Results.

The two B-readers were asked to select representative regions of interest (ROIs) for the 124 subjects in this part of the study. The pulmonologists identified 963 ROIs where each ROI represented one of four categories: Normal tissue, parenchymal abnormalities, blood vessels, and ribs. Table 14 shows the distribution of number of ROIs for each category. It also presents the distribution of type of abnormalities and their severity. The B-readers worked together and reached a consensus on each of the ROIs. Using a software tool developed specifically for this purpose, the pulmonologists would draw a rectangular box to encompass the region which they wished to sample. The pulmonologists could adjust the dimensions of the box (ROI) to any size. The computer program would automatically save the location of the ROI, its size, and its classification, according to the options in Table 14. A data base was created which identified the chest radiograph by a coded subject number. The entire, uncompressed chest radiograph was saved along with its ROI information. CDs were written to archive the data. Figure 6 shows an example of the computer interface used by the readers to read and document the 124 chest radiographs.

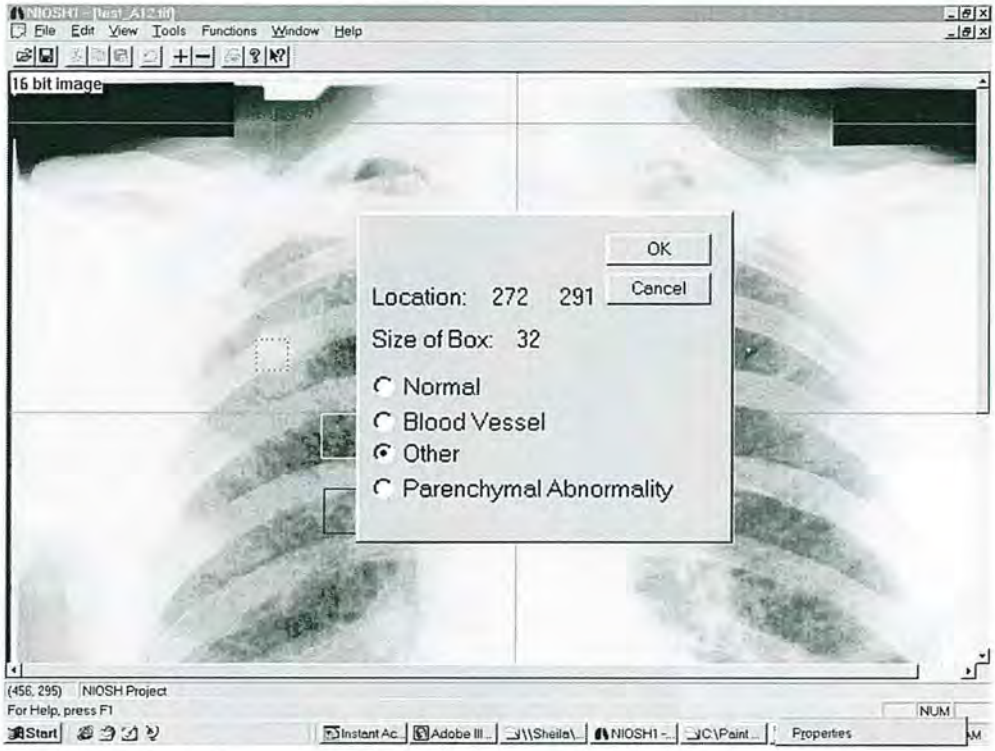

Figure 6. CARRS computer screen depicting window with chest radiograph and dialogue box with class selection for ROI.

Table 14. Summary of ROIs selected and "read" by the two B-readers.

TOTAL	Normal	Parench.	Vessel	Ribs
963	170	83	556	154
100%	18%	9%	58%	16%
			Total ROIs	963
Type ↓ & Severity→ (Total by type)		**1**	**2**	**3**
p	(3)	3	0	0
q	(69)	56	12	1
r	(2)	0	1	1
s	(0)	0	0	0
t	(8)	3	2	3
u	(1)	1	0	0

The statistical (co-occurance) features were calculated for each of the ROIs. Using the co-occurance statistics for the ROI as the feature vector, LAPART was trained on 120 of the 124 subjects. The LAPART creates a set of patterns that "represent" each of the classes, *i.e.* normal, abnormal, blood vessels, and ribs. The four subjects that were not part of the training were used to test the performance of the classifier. The four subjects were selected randomly, but constrained to include at least one subject with a parenchymal abnormality. This ensured that the classifier had at least one example of the abnormal ROI to test. Table 15 shows the results from each of the ten LAPART training and testing sessions.

Along the left side of each table is the number of ROIs in each category as established by the "ground truth" (**GT**) process described earlier. The classification made by the LAPART for each ROI is given in

the matrix. For example, all ground truth normals were correctly classified (see first row first column under normal in Test #1). One abnormal was classified by LAPART as blood vessels (see second row under vessel in Test #1).

Table 15. Results from the 10 LAPART tests.

Ground Truth:	LAPART	Test # 1 Normal	Abnormal	Vessel	Rib, etc.
6	Normal	6	0	0	0
4	Abnormal	0	3	1	0
20	Vessel	0	0	20	0
7	Rib, etc.	0	0	0	7
37		6	3	21	7

Ground Truth:	LAPART	Test # 2 Normal	Abnormal	Vessel	Rib, etc.
7	Normal	6	1	0	0
4	Abnormal	0	3	1	0
21	Vessel	0	0	21	0
6	Rib, etc.	0	0	0	6
38		6	4	22	6

Ground Truth:	LAPART	Test # 3 Normal	Abnormal	Vessel	Rib, etc.
6	Normal	6	0	0	0
3	Abnormal	0	3	0	0
23	Vessel	0	0	23	0
7	Rib, etc.	0	0	0	7
39		6	3	23	7

Ground Truth:	LAPART	Test # 4 Normal	Abnormal	Vessel	Rib, etc.
6	Normal	5	1	0	0
3	Abnormal	0	3	0	0
19	Vessel	0	0	19	0
7	Rib, etc.	0	0	0	7
35		5	4	19	7

Ground Truth:	LAPART	Test # 5 Normal	Abnormal	Vessel	Rib, etc.
5	Normal	5	0	0	0
4	Abnormal	0	3	1	0
18	Vessel	0	0	18	0
5	Rib, etc.	0	0	0	5
32		5	3	19	5

Ground Truth:	LAPART	Test # 6 Normal	Abnormal	Vessel	Rib, etc.
4	Normal	4	0	0	0
5	Abnormal	0	4	1	0
22	Vessel	0	0	22	0
5	Rib, etc.	0	0	0	5
36		4	4	23	0 5

Ground Truth:	LAPART	Test # 7 Normal	Abnormal	Vessel	Rib, etc.
5	Normal	5	0	0	0
3	Abnormal	0	3	0	0
18	Vessel	0	0	18	0
7	Rib, etc.	0	0	0	7
33		5	3	18	7

Ground Truth:	LAPART	Test # 8 Normal	Abnormal	Vessel	Rib, etc.
7	Normal	7	0	0	0
3	Abnormal	0	3	0	0
22	Vessel	0	0	22	0
6	Rib, etc.	0	0	0	6
38		7	3	22	6

Ground Truth:	LAPART	Test # 9 Normal	Abnormal	Vessel	Rib, etc.
6	Normal	6	0	0	0
3	Abnormal	0	3	0	0
18	Vessel	0	0	18	0
7	Rib, etc.	0	0	0	7
34		6	3	18	7

Ground Truth:	LAPART	Test # 10 Normal	Abnormal	Vessel	Rib, etc.
5	Normal	5	0	0	0
3	Abnormal	0	3	0	0
17	Vessel	0	0	17	0
6	Rib, etc.	0	0	0	6
31		5	3	17	6

A summary of the 10 tests is given in Table 16. The vigilance setting for running the LAPART was set to 0.7. Training required approximately 15 seconds on a personal computer with a Pentium (400 MHz) processor.

Table 16. Summary of LAPART tests.

GT Total		Normal	Abnormal	Vessel	Rib, etc.
57	Normal	**55**	**2**	**0**	**0**
35	Abnormal	**0**	**31**	**4**	**0**
198	Vessel	**0**	**0**	**198**	**0**
63	Rib, etc.	**0**	**0**	**0**	**63**

Sensitivity, specificity, and accuracy are defined as follows:

Sensitivity = TP / (TP + FN)

Specificity = TN / (TN + FP)

Accuracy = (TP + TN)/Total

Table 17 combines the normal class with blood vessels and rib ROIs to form a single "normal" category. For sensitivity, specificity and sensitivity calcualtions, the true positive, false negative, true negative, and false positive were defined by combining the blood vessel and rib category with the normal ROIs into a single "normal" class. Then the truth table for the 10 tests is given below. The LAPART classification is given in the columns and the B-reader ground truth is in the rows. For example, the B-readers classified 322 normals and 35 parenchymal abnormal ROIs. The LAPART classified the ROIs as 324 normal and 33 abnormals. The sensitivity was calculated as 0.886 with a specificity of 0.994. The total accuracy was 0.983.

Table 17. Normal (including blood vessels and ribs) versus ROIs presenting with parenchymal abnormalities.

	All Normal	Parench.
All Normal	320	2
Parench.	4	31

A series of experiments using LAPART to train only on ROIs that had been classified by the B-readers as normal or presenting with parenchymal abnormalities were used in the training and testing. All ROIs for ribs and vessels eliminated from the training and testing. The summary of ten LAPART sessions is presented in Table 18.

Table 18. Training on ROIs labeled as normal or as presenting parachymal abnormalities.

Summary:

		Normal	Parench.	Vessel	Rib, etc.
50	Normal	**48**	**2**	**0**	**0**
30	Parench.	**0**	**30**	**0**	**0**
0	Vessel	**0**	**0**	**0**	**0**
0	Rib, etc.	**0**	**0**	**0**	**0**

The sensitivity improved to 1.00 and the specificity was reduced to 0.96.

A series of tests were conducted using only the ROIs presenting with a parenchymal abnormality. There were 68 ROIs that were read as type "q" and severity 1 or 2. These 68 ROIs were used to train and test the LAPART's ability to separate the two classes of severity for type "q". 18% percent of the ROIs (12-

14 per test) were withheld from the training and used for testing. The tests were constrained so that at least one ROI with a severity of 2 was withheld from the training and used for testing. The results of the 10 tests are presented in Table 19.

Table 19. LAPART results for training to distinguish severity levels.

	LAPART Test # -->		1	
GT	Severity :	1	2	3
12	1	11	1	0
2	2	0	2	0
0	3	0	0	0
14	14	11	3	0

	LAPART Test # -->		2	
GT		1	2	3
12	1	11	1	0
2	2	0	2	0
0	3	0	0	0
14	14	11	3	0

	LAPART Test # -->		3	
GT		1	2	3
10	1	9	1	0
2	2	0	2	0
0	3	0	0	0
12	12	9	3	0

	LAPART Test # -->		4	
GT		1	2	3
9	1	9	0	0
2	2	0	2	0
0	3	0	0	0
11	11	9	2	0

	LAPART Test # -->		5	
GT		1	2	3
11	1	10	1	0
4	2	1	3	0
0	3	0	0	0
15	15	11	4	0

	LAPART Test # -->		6	
GT		1	2	3
9	1	9	0	0
3	2	0	3	0
0	3	0	0	0
0	0	0	0	0
12	12	9	3	0

	LAPART Test # -->		7	
GT		1	2	3
8	1	8	0	0
3	2	1	2	0
0	3	0	0	0
11	11	9	2	0

	LAPART Test # -->		8	
GT		1	2	3
8	1	7	1	0
2	2	0	2	0
0	3	0	0	0
10	10	7	3	0

	LAPART Test # -->		9	
GT		1	2	3
8	1	7	1	0
3	2	0	3	0
0	3	0	0	0
11	11	7	4	0

	LAPART Test # -->		10	
GT		1	2	3
9	1	8	1	0
2	2	0	2	0
0	3	0	0	0
0	0	0	0	0
11	11	8	3	0

The summary of the 10 tests is given in Table 20. Severity level 3 ROIs were excluded from the tests. There were only 5 ROIs that were read as severity 3. It was felt that the sample size would be too small to allow the neural network to generalize sufficiently the feature characteristics of this class in order to establish a class feature pattern. The sensitivity was 0.920. The specificity is 0.927 and the accuracy 0.926.

Table 20. Summary of tests for classifying severity levels.

GT Total		1	2	3
96	1	89	7	0
25	2	2	23	0
0	3	0	0	0

E. Discussion and Conclusions.

Not surprisingly, the features describing the normal tree-like structure of the vascularity in the normal tissue in a few cases resembled some of the features derived from opacities and profusions presented in the example of parchymal abnormalities, which resulted in the six misclassifications in Table 16. Two ROIs that are representative of normal tissue were classified as abnormal and four ROIs presenting with parenchymal abnormalities were classified by the neural network as blood vessels. These misclassifications demonstrate the problem faced by pulmonologists in trying to detect and differentiate pneumoconioses opacities from the normal pulmonary vascularity. The problem is often exacerbated by the partial masking of the vacularity by the opacities.

Nevertheless, the sensitivity, specificity and accuracy demonstrated by the combined techniques of texture analysis and feature vector classification using the LAPART neural network indicate a high potential for computer-aided diagnosis. From Table 16 one can see that 98.3 % of the ROIs were correctly classified by tissue type and abnormality. The shows that the classifier does very well in differentiating the four classes of ROIs. A future extension of the program would search out or sample different regions of the x-ray and produce a classification as to the type of ROI, *i.e.* tissue type. A second classifier (or the process could be combined) would further segment the classification of the abnormal ROIs into a severity level. This potential is documented the in last set of experiments. Table 20 shows the performance of the neural network in classifying the severity levels of a type "q" abnormality.

The amount of data samples for the different types and severity limited the study to only the "q" type. The results of the test on 121 randomly selected ROIs shows sensitivities and specificities in excess of 0.90. These results are consistent with those of others who have used statistical methods for characterizing texture in chest radiographs. Showing that statistical methods, such as the use of the co-occurrence statistics, produce positive results for the classification of abnormalities in chest radiographs was a necessary finding in order to progress to the testing and implementation of a commercial computer-aided chest radiograph diagnostic system.

Because this was a short-term feasibility study, the available database was necessarily limited. Large data bases consisting of tens of thousands of chest radiographs from miners in the Southwestern United States are available and will be accessed in Phase II.

The research and development of the computer-aided chest radiograph diagnostic system is embodied in the software analysis system and computer interface that has been developed to incorporate the manipulation of chest radiographs to select ROIs and produce the classification using the neural network. The current computer-aided chest radiograph diagnostic system is a prototype suitable for much of the analysis and testing that must continue in Phase II with a more extensive database. A process will be described in the Phase II proposal whereby the tedious task of ROI labeling will be progressively more automatic as the system evolves. Growth of the computer-aided chest radiograph diagnostic system will be evolutionary, meaning that manually intensive operations will become more automatic, while always under the control of the pulmonologist.

One of the advantages of the methods adopted for the development and demonstration of the computer-aided chest radiograph diagnostic system is the ability for it to grow to include large databases without becoming computationally limited. And this attribute stems primarily from the type of classifier that has been developed and tested. The laterally primed adaptive resonance theory neural network is extremely fast, even on standard personal computers. It has been tested as an image segmentation algorithm where it processes millions of feature vectors in less than 2 minutes. Its architecture allows one to continue to update its learned patterns with new data without excessive retraining times needed for many neural networks. Another significant attribute lies in its presentation of "human-friendly" rules rather than complex, un-interpretable weights given by many neural networks such as back-propagation.

Results show great promise for a Phase II. The purpose of the Phase I research was "to demonstrate feasibility." This has been accomplished.

F. References.

1. Parker, D.L., A.B. Bender, S. Hankinson, and D. Aeppli, *Public health implications of the variability in the interpretation of "B" reading for pleural changes,* J. Occup. Med, 1988(31): p. 775-780.
2. Haralick, R.M. *Statistical and structural approaches to texture*, in *IEEE.*
3. Healy, M.J. and T.P. Caudell. *Acquiring rule sets as a product of learning in a logical neural architecture*, in *IEEE Trans. Neural Networks*. 1997.
4. Carpenter, G.A. and S. Grossberg, *A massively parallel architecture for a self-organizing neural pattern recognition machine,* Computer Vision, Graphics and Image Processing, 1987. 37: p. 54-115.
5. *Cancer Facts and Figures - 1966*, American Cancer Society: Atlanta, GA.
6. *Reducing the Health Consequences of Smoking - 25 Years of Progress: A Report of the Sergion General*, 1989, U. S. Department of Health and Human Services, Office on Smoking and Health: Rockville, MD.
7. Page, H.S. and A.J. Asire, *Cancer Rates and Risks*, 1985, National Cancer Institute: Bethesda, MD.
8. *Advance report of final mortality statistics, 1990*, 1993, National Center for Health Statistics, Month. Vital Stat. Rep. p. 41(7) (suppl) DHHS Publication 93-1120.
9. Levin, L. and D.C. Levin, *Morbidity and quality of life in patients with COPD*, in *Clinical Epidemiology of Chronic Obstructive Pulmonary Disease*, M.J. Hensley and N.A. Saunders, Editors. 1989, Marcel Dekker: New York, NY. p. 45-54.
10. Ridlich, C.A., *Pulmonary fibrosis and interstitial lung diseases*, in *Occupational and Environmental Respiratory Disease*, P. Harber, M.B. Schenker, and J.R. Balmes, Editors. 1996, Mosby-Year Book, Inc.: St Louis, MO, p. 2116-227.
11. Tomatis, L. and Ed. *Cancer: Causes, Occurrence and Control,* in *International Agency for Research on Cancer*. 1990, Lyon, France.
12. Crystal, R.G., *Interstitial lung disease*, in *Textbook of Medicine*, J.B. Wyngaarden, L.H. Smith, and J. Bennett, Editors, 1992, W. B. Saunders: Philadelphia, PA.
13. Coultas, C.B., R.E. Zumwalt, W.C. Black, and R.E. Sabonya, *The epidemiology of interstitial lung diseases,* Am J Respir Crit Care Med, 1994. 150: p. 967-972.
14. Epler, G.R., *Chest films: "Underused tool in interstitial disease?",* J Respir Dis, 1997(18): p. 168-180.
15. Goldring, J.M., D.S. James, and H.A. Anderson, *Chronic lung diseases*, in *Chronic Disease Epidemiology and Control*, R.C. Brownson, P.L. Remington, and J.R. Davis, Editors. 1993, American Public Health Assoc.: Washington, DC. p. 169-197.
16. Harber, P., *Surveillance and screening of respiratory disease*, in *Occupational and Environmental Respiratory Disease*, P. Harber, M.B. Schenker, and J.R. Balmes, Editors, 1996, Mosby-Year Book, Inc.: St. Louis, MO. p. 216-227.
17. *Guidelines for the use of ILO international classification of radiographs of pneumoconioses*, 1980, International Labour Office: Geneva.
18. Shipley, R.T., *The 1980 classification of radiographs of the pneumoconioses*, in *Occupational Lung Disease*, 1992, Radiologic Clinics of North America. p. 135-1145.
19. McCloud, T.C., C.B. Crrington, and E.A. Gaensler, *Diffuse infiltrative lung disease: A new scheme for description,* Radiology, 1983(149): p. 353-363.
20. David, A., *Present use and trends in the development of the ILO international classification of radiographs of pneumoconioses.* in *The VIIth International Pneumoconiosis Conference*, 1988. Pittsburgh.
21. David, A., *X-ray films and asbestosis,* Chest, 1989. 96: p. 952.
22. Ducatman, A.M., W. Yang, and S.A. Forman, *B-readers' and asbestos medical surveillance,* J. Occup. Med, 1988(30): p. 644-647.
23. Soliz, P. and P. Investigator). *Semi-automatic, quantitative analysis of human fundus.* in *NEI Contract #1 R43 ER10951-01, 1995 and NEI Contract #1 R43 EY10951-03, 1997.*
24. Amandus, H.E., E.P. Pendergrass, and W.K.C. Morgan, *Pneumoconiosis: inter reader variability in the classification of a type of small opacities in the chest radiograms,* AJR, 1974(12): p. 740-743.

25. Bourbeau, J. and P. Enst, *Between- and within-reader variability in the assessment of pleural abnormality using the ILO 1980 International Classification of Pneumoconiotic chest film,* Occup. Med., 1988(14): p. 537-543.
26. Muir, D.C.F., C.D. Bernholz, W.K.C. Morgan, J.O. Roos, J. Chan, W. Maehle, J.A. Julian, and A. Sebestyen, *Classification of chest radiographs for pneumoconiosis: a comparison of two methods of reading,* British Journal of Industrial Medicine, 1992(49): p. 869-871.
27. Kruger, R.P., W.B. Thompson, and A.F. Turner. *Computer Diagnosis of Pneumoconiosis.* in *IEEE Transactions on Systems, Man, and Cybernetics*, January 1974.
28. Ledley, R.S., H.K. Huang, and L.S. Rotolo, *A Texture Analysis Method in Classificatoin of Coal Workers' Pneumoconiosis,* Comput. Biol. Med., 1974. 5: p. 53-67.
29. Monnier-Cholley, L., H. MacMahon, S. Katsuragawa, J. Morishita, T. Ishida, and K. Doi, *Computer-Aided Diagnosis for Detection of Interstitial Opacities on Chest Radiographs,* AJR, December 1998. 171: p. 1651-1656.
30. Haralick, R.M., K.S. Shammigam, and I. Dinstein. *Textural features for image classification.* in *IEEE Transactions on Systems, Man, and Cybernetics*, November 1973.
31. Katsuragawa, S., K. Doi, and H. MacMahon, *Image feature analysis and computer-aided diagnosis in digital radiography: Detection and characterization of interstitial lung disease in digital chest radiographs,* Am. Assoc. Phys. Med., 1988. 15(3): p. 311-319.
32. Kim, J.K. and H.W. Park. *Statistical Textural Featured for Detection of Microcalcifications in Digitized Mammograms*, in *IEEE Transactions on Medical Imaging*. March 1999.
33. Penedo, M.G., M.J. Carreira, A. Mosquera, and D. Caballo, *Computer-Aided Diagnosis: A Neural-Network-Based Approach to Lung Nodule Detection*, in *IEEE Transactions on Medical Imaging*, December 1998.
34. Lo, S.B., S.A. Lou, J. Lin, M.T. Freedman, M.V. Chien, and S.K. Mun, *Artificial convolution neural network techniques and applications for lung module detection,* Transactions on Medical Imaging, 1995, 14(4): p. 711-718.
35. Fukushima, K., S. Miyake, and T. Ito, *Neocognitron: A neural network model for a mechanism of visual pattern recognition*, in *IEEE Trans. Sys. Man, and Cyber,* 1983.
36. Schnorrenberg, F., C.S. Pattichis, K.C. Kyriacou, and C.N. Schizas, *Computer-Aided Detection of Breast Cancer Nuclei*, in *IEEE Transactions of Information Technology in Biomedical*, June 1997.
37. Savol, A.M., C.C. Li, and R.J. Hoy. *Computer-Aided Recognition of Small Rounded Pneumoconiosis Opacities in Chest X-Rays.* in *IEEE Transactions on Pattern Analysis ;and Machine Intelligence*, 1980.
38. Giger, Dio, and H. MacMahon, *Image feature analysis and computer-aided diagnosis in digital radiography. Automated detection of nodules in peripheral lung fields,* Am. Assoc. Phys. Med., Mar/Apr 1988, 15(2): p. 158-166.
39. Soliz, P. and G.W. Donohoe, *Adaptive Resonance Theory Net for Funds Image Labeling*, in *World Congress on Neural Networks*. September, 1996, San Diego, CA.
40. Donohoe, G.W., P. Soliz, and S.C. Nemeth, *Computer Aided Image Analysis for Background Diabetic Retinopathy.* in *SPIE Medical Imaging Symposium on Medical Imaging*, Feb. 1998, San Diego, CA.
41. Farnath, D., S.C. Nemeth, P. Soliz, and A. Das, *Analysis of retinal neovascularization with a computer-assisted digital fundus image diagnostic system*, in *Association for Research in Vision and Ophthalmology (ARVO)*, May, 1998, Ft. Lauderdale, FL.
42. Fleiss, J.L., J. Cohen, and B.S. Everitt, *Large sample standard errors of kappa and weighted kappa,* Psychological Bulletin, 1969, 72(5): p. 323-327.
43. Aattfield, M.D. and G.R. Wagner, *A report on a workshop on the National Institute for Occupational Safety and Health B Reader Certification Program,* JOM, 1992(34): p. 875-878.
44. Unser, M. *Sum and difference histograms for textrure classifcation*, in *IEEE Trans. Patterns Anal. Machine Intelligence*, 1986.
45. Laws, K.I. *Texture Energy Measures.* in *Proc. Image Understanding Workshop*. 1995.
46. Mandelbrot, B.B., *Fractals: Form, Chance and Dimension*, 1977, San Francisco, CA: Freeman. 468.
47. Norton, A., *Generation and display of geometric fractals in 3-D.* Comput. Graphics, 1982(16(3)): p. 61-67.
48. Kawaguchi, Y., *A morphological study of the form of nature,* Computer Graphics, 1982(16(3)): p. 223-232.

49. Witkin, A.P., *Recovering surface shape and orientation from texture.* Artificial Intelligence, 1981(17): p. 17-47.
50. Keller, S.C. and R.M. Crownover, *Texture description and segmentation through fractal geometry.* Comput. Vision Graph. Image Process, 1989(45): p. 150-166.
51. Kaizer, H.A., *A quantification of textures on aerial photographs*, . 1955, Boston University: Boston, MA.
52. Dyer, C.R. and A. Rosenfeld. *Fourier texture features: Suppression of aperature effects.* in *IEEE Trans. Syst., Man., Cybern.* 1976.
53. Pollack, I., *Visual discrination thresholds for one- and two-dimensional Markov spatial constraints.* Perception Psycholophys, 1972. 12(2A): p. 161-167.
54. Manlebrot, B., *The Fractal Geometry of Nature.* 1982, New York: W. H. Freeman.
55. Julesz, B. *Visual pattern discrimination.* in *IEEE Trans. Inf. Theory.* 1962.
56. Getty, D.J., C.J. Pickett, C.J. D'Orsi, and J.A. Swets, *Enhanced interpretation of diagnostic images.* Investigative Radiology, 1988. 23: p. 240-252.
57. Huo, Z., M.L. Giger, C.J. Vyborny, R.A. Schimidt, K. Doi, and P. Lu, *Computer-aided diagnosis in mammography: Classification of malignant and benign masses.* Radiology, 1993. 189: p. 318.
58. Caudell, T.P. and S.D.G. Smith. *A data compressed ART-1 Neural Network Algorithm.* in *International Joint Conference for Neural Networks.* July 1991. Seattle, WA.
59. Grossberg, S. *The Adaptive Brain I: Cognition, Learning Reinforcement, and Rhythm, and the Adaptive Bran II: Vision, Speech, Language, and Motor Control.* 1986. Elsevier/North-Holland, Amsterdam.
60. Healy, M.J. and T.P. Caudell. *Discrete stack interval representation and Fuzzy ART1.* in *World Congress on Neural Networks.* July 1993. Portland, OR.
61. Soliz, P., *Digital Fundus Image Diagnostic System.* NEI Contract #1 R43 EY10951-01 1995 and NEI Contract #1 R43 EY10951-03 1997-1998, .

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Public Health Service
Centers for Disease Control
and Prevention (CDC)

Memorandum

Date: March 23, 2001

From: Roy M. Fleming, Sc.D., Director, Research Grants Program RMF
Office of Extramural Programs, NIOSH, D30

Subject: Final Report Submitted for Entry into NTIS for Grant 1 R43 OH003595-01.

To: William D. Bennett
Data Systems Team, Information Resources Branch, EID, NIOSH, P03/C18

The attached final report has been received from the principal investigator on the subject NIOSH grant. If this document is forwarded to the National Technical Information Service, please let us know when a document number is known so that we can inform anyone who inquires about this final report.

Any publications that are included with this report are highlighted on the list below.

Attachment
cc: Sherri Diana, EID, P03/C13

List of Publications - None

Title: Computer-Assisted Chest Radiograph Reader
Investigator: Peter Soliz, Ph.D.
Affiliation: Kestrel Corporation
City & State: Albuquerque, NM
Telephone: (505) 345-2327
Award Number: 1 R43 OH003595-01
Start & End Date: 9/30/1998–3/31/1999
Total Project Cost: $99,320
Program Area: Not NORA
Key Words:

Abstract:
To reduce inter- and intra-reader variability in diagnosing chest radiographs, a neural network-based computer-aided diagnostic system was developed and tested. The results of an experiment with 124 digitized chest radiographs, demonstrated high degrees of sensitivity and specificity in classifying chest radiographs. The use of a computer-assisted chest radiograph reader eliminated the inconsistencies in the human readers. The Computer-assisted Chest Radiograph Reader System (CARRS) applies recognized principles in the psychophysics of human vision, incorporates neural network-based image analysis and integrates these with a graphical user interface. Advances in digital image processing, and classification techniques have made CARRS feasible for meeting screening, research and development, and clinical requirements.

Through the adoption of the International Labor Organization (ILO) classification procedures, it had been hoped that reader variation in the classification of parenchymal abnormalities could be minimized. The ILO classification of the pneumoconioses is based on a structured procedure for detecting and characterizing patterns on chest radiographs. Numerous studies have shown, however, that inter- and intra-observer variability of radiograph readings by trained medical personnel has persisted.

The methodology was implemented through the following tasks: 1) From a data base of several thousand patients, a set of 205 chest radiographs were manually graded by two pulmonologists; 124 of the films were then digitized at 12-bit high spatial resolution. 2) Textural features were calculated using high order statistical techniques. The features were classified by the pulmonologists to "train a neural network to extract classification rules chest radiographs based on the ILO methodology. 3) The neural network classification from the graded system was tested using 65 chest radiographs.

For 5-10 areas selected by the pulmonologist on the chest radiograph, a feature vector composed of image characteristics such as density distribution, entropy, fractal dimension, opacity counts, shape, etc. was calculated. This feature vector characterized numerically the areas used by the pulmonologist to grade the radiograph. The neural network trained on the same regions used by the pulmonologist, and through a quantitative feature vector, "learned" the characteristics of each ILO classification.

The laterally-primed adaptive resonance theory (LAPART) neural network was selected. LAPART presents the results of the training in human interpretable "rules." LAPART trains in a single pass ("fast") making it attractive for a clinical setting. To demonstrate that calculated characteristics of chest radiographs could be used to train the LAPART to classify the radiographs using objective and quantitative parameters, a pilot study was conducted. 32 statistical features from 124 digitized chest radiographs were calculated. Statistical parameters were selected based on their contribution to the separation of classes. The parameters included entropy, contrast, fractal dimension, and co-occurrence statistics. The computational efficiency of LAPART was demonstrated in this experiment. LAPART was trained on a Pentium 200. Training time for 10 experiments, where the training and testing radiographs were randomly selected, averaged 15 seconds per experiment.

In conclusion these experiments showed that LAPART is a worthy candidate for basing a semi-automatic chest radiograph classification system. The experiments demonstrated that CARRS has low classification variability and has a significantly high accuracy. The results of each of the 10 experiments showed that the LAPART neural network algorithm could be trained to "learn" the extent of the "hypercubes" which represented each of the six classifications in 32-dimensional space. LAPART classification accuracy averaged 98 %. "Truth" was determined by the two radiologists.

Publications
No publications to date.