A Screening System for the Assessment of Opacity Profusion in Chest Radiographs of Miners with Pneumoconiosis

M.S. Pattichis [1], C.S. Pattichis[1,2], C.I. Christodoulou [2], D. James[3,4], L. Ketai[3], P. Soliz[5]
[1] Department of Electrical and Computer Engineering, University of New Mexico, NM
[2] Department of Computer Science, University of Cyprus, Cyprus
[3] University of New Mexico Health Sciences Center, Albuquerque, NM
[4]Miners Colfax Medical Center, Raton, NM
[5]Kestrel Corporation, Albuquerque; NM
pattichis@eece.unm.edu, *costas@eece.unm.edu*, *cschr2@ucy.ac.cy,*
dsjames@salud.unm.edu, lketai@salud.unm.edu, *zilospx@rt66.com*

ABSTRACT

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

I. Introduction

Individuals who have been exposed through their occupations to high levels of dust, asbestos, or other particulates are at risk for interstitial lung diseases such as pneumoconiosis. These individuals must have periodic examinations including chest radiographs to monitor for signs of the opacities associated with pneumoconiosis [1]. This study focused on a cohort of individuals whose occupation resulted in exposure to coal mining related dust. The International Labor Organization (ILO) has established an objective protocol to score the degree of profusion, size, and shape of opacities [2]. The protocol depends on the experience and perception of the radiologist to compare the chest radiograph with a set of standards and make a decision as to the profusion of opacities. This perception-based approach results in inter- and intra-individual variability [3]. The inter- and intra-individual variability is especially problematic for the low profusions, e.g. 0/0 (normal), 0/1, and 1/0.

Interpreting chest radiographs with fine degrees of precision, as is called for in the ILO protocol, presents a significant challenge, even to the most experienced physician. The chest radiograph has a number of anatomical structures that are superimposed on the areas of interest. For example, the normal presentation of blood vessels can be misinterpreted as small opacities that indicate interstitial lung disease. The demand for precision that is near the limits of the human perceptual system and the confounding structures in the chest radiograph suggests solutions that may be offered by computer algorithms. In particular, algorithms that characterize the fine textural features of the ribs and inter-rib tissue, the parenchyma and use these features as a basis for classify profusion and opacity size and shape are discussed in this paper.

The aim of this study is to develop a semi-automated screening system based on texture analysis for the assessment of opacity profusion in chest radiographs of miners with pneumoconiosis. The system is currently semi-automated in that the physicians are asked to select circular Regions Of Interest (ROIs) that are then used for training, and consequently testing the overall system.

Fifth IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI'02)
0-7695-1537-1/02 $17.00 © 2002 IEEE

The fundamental contribution of this paper is to demonstrate a system that can distinguish between normal and abnormal Regions of Interest. An automated classification system using a Self-Organizing Map (SOM), the backpropagation (BP), and the radial basis function (RBF) neural networks classifiers, as well as the statistical k-nearest neighbour (KNN) classifier are used for classifying the feature vectors over each Region of Interest into two classes: profusion 0 and profusion 1(q and r).

II. Material

Chest radiographs were provided by the Miners Colfax Medical Center, Raton, NM. The radiographs were of coal-mine or silica dust exposed miners participating in a health screening program. The radiographs were digitized at 300 dpi and 12 bits resolution. Two certified B-readers were asked to provide ground truth by selecting representative regions of interest (ROI) in the radiographs. By consensus, each ROI was given a profusion and opacity size. Profusion is given in twelve categories, from normal (0/0), 0/1, 1/0, 1/1, etc to 3/+. Size of the rounded opacities is given as p (0-1.5mm), q (1.5-3.0mm), and r (3.0-10.0mm). A total of 236 ROIs (166, 49, and 21 with profusions of category (shape and size) 0, 1(q), and 1(r), respectively) were classified from 74 digitized chest radiographs by two B-readers. The most inter- and intra-reader variability is in the low profusions, such as those between 0 and 1. Most of the opacity sizes found in the radiographs from the Miners Colfax Medical Center database are q. Two sets of data were selected for developing the screening system: 1) for training the system, and 2) for evaluating its performance. For training the system 50 ROIs of profusion 0 and 50 ROIs of profusion 1(q and r) were used, whereas for evaluation of the system the remaining 116 and 20 ROIs were used respectively.

III. Feature Extraction

In this study, the following texture features were extracted from the ROIs:

A. Spatial Gray Level Dependence Matrices (SGLDM).

The spatial gray level dependence matrices as proposed by Haralick et al. [4] are based on the estimation of the second-order joint conditional probability density functions that two pixels (k, l) and (m, n) with distance d in direction specified by the angle χ, have intensities of gray level i and gray level j. Based on the probability density functions the following texture measures [4] were computed: 1) Angular second moment, 2) Contrast, 3) Correlation, 4) Sum of squares: variance, 5) Inverse difference moment, 6) Sum average, 7) Sum variance, 8) Sum entropy, 9) Entropy, 10) Difference variance, 11) Difference entropy, and 12), 13) Information measures of correlation. For a chosen distance d (in this work d=1 was used) and for angles $\chi = 0^o$, 45^o, 90^o and 135^o, we computed four values for each of the above 13 texture measures. In this work, the range of these four values were computed for each feature.

B. Gray Level Difference Statistics (GLDS).

The GLDS algorithm [5] uses first order statistics of local property values based on absolute differences between pairs of gray levels or of average gray levels in order to extract the following texture measures: 1) Contrast, 2) Angular second moment, 3) Entropy, and 4) Mean.

IV. Screening System

The self-organizing map (SOM) classifier [6], the backpropagation (BP) [7], and the radial basis function (RBF) [7] neural networks classifiers, as well as the statistical k-nearest neighbour (KNN) [8] classifiers were used for the screening system. The system was trained to classify two classes: profusion 0 and profusion 1(q and r). The SGLDM and GLDS features sets were used as input to the classifier as described in the previous section. All features were standardised by subtracting the mean and dividing by the standard deviation for the whole data set.

V. Results

For each feature parameter, the results are summarized in Table I. For each textural feature, the resulting probability density function (pdf) is given in terms of the median and the inter quartile range (iqr). For the iqr, the difference between the 75th and the 25th percentiles is taken as a measure of the spread of each pdf.

Next, we considered the discriminating power of each individual textural feature for differentiating between normal and abnormal ROIs. The Wilcoxon rank-sum test was used as a non-parametric hypothesis test whether there are significant differences (at a significance level a=0.05) between 0 vs 1(q), and 0 vs 1(r), for 14 and 12 features respectively. A value of 1 in last column in the table indicates that significant differences were found, while a value of 0 indicates that no significant differences were found. Overall, the results were mostly positive in that most individual features showed significant differences between normal and abnormal ROIs.

Table I Median (Med) and inter quartile range (iqr) for the SGLDM and GLDS feature sets for classes 0, 1(q) and 1(r). Results of non-parametric Wilcoxon rank sum test between profusion category 0 versus that of profusions 1(q) and 1(r). ´1´ indicates significant difference, and ´0´ no significant difference at a=0.05.

FEATURE SET	Class 0		Class 1(q)		Class 1(r)		Wilcoxon test.	
	Med	iqr	Med	iqr	Med	iqr	0 vs 1(q)	0 vs 1(r)
SGLDM								
Angular second moment	0.0004	4E-04	0.0003	0.0003	3E-04	0.0002	0	0
Contrast	48.882	89.99	104.87	238.34	161.8	250.08	1	1
Correlation	0.0262	0.028	0.0304	0.0499	0.019	0.0289	0	0
Sum of squares: variance	7.453	23.05	14.196	36.634	23.68	53.731	1	0
Inverse differ. Moment	0.0882	0.031	0.1105	0.036	0.109	0.0381	1	1
Sum average	0.205	0.251	0.3848	0.4684	0.605	0.7674	1	0
Sum variance	73.524	161.7	169.82	335.24	257.2	442.24	1	1
Sum entropy	0.0156	0.014	0.0286	0.0189	0.018	0.0193	1	1
Entropy	0.4358	0.253	0.5795	0.3121	0.659	0.3476	1	1
Difference variance	16.903	30.71	31.912	60.547	49.49	73.751	1	1
Difference entropy	0.4683	0.277	0.638	0.343	0.689	0.44	1	1
Information measures	0.0939	0.04	0.1219	0.0503	0.123	0.0407	1	1
of correlation	0.0286	0.028	0.0309	0.0389	0.021	0.0176	0	0
GLDS								
Contrast	248.47	388	459.39	856.48	607.5	887.62	1	1
Angular second moment	0.0375	0.024	0.0268	0.0252	0.023	0.0257	1	1
Entropy	3.4705	0.672	3.7827	0.8057	3.914	0.8687	1	1
Mean	12.041	8.625	17.077	15.387	19.35	14.253	1	1

Figure 1 illustrates the SOM map for the training set for the SGLDM feature set. The high degree of overlap between the boundaries of the two classes is shown.

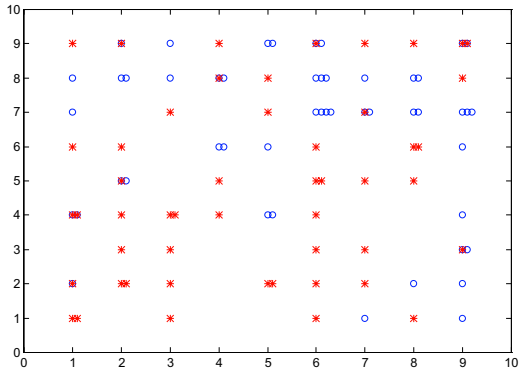

Fig. 1 Distribution of 100 ROIs of the training set (50 and 50 with profusion 0 and 1(q and r) respectively) on a 9x9 SOM using as input the SGLDM feature set (o = profusion 0, * = profusion 1(q and r)). Similar ROIs are assigned to neighbouring output matrix nodes.

Table II tabulates the percentage of correct classifications for the SGLDM, and GLDS feature sets for the evaluation set (i.e. 116 cases class 0 and 20 cases class 1) using the SOM, BP, RBF, and KNN classifiers. The highest percentage of correct classifications score for the evaluation set were 75%, 72% and 71% for the SGLDM feature set for the BP, KNN and SOM classifiers respectively. The BP was trained with one hidden layer of 20 neurons, whereas the RBF was trained with 97 neurons. Both the BP and RBF were trained for 50 epochs, further increase in the number of epochs caused a drop in the correct classifications score. It is very clearly shown that the SGLDM feature set gave better results than the GLDS feature set. For the SOM in order to derive a more reliable estimate of the ROI class, a neighbourhood window centered at the winning node taking into consideration the majority of the labels in the window region was applied [9].

Table II Percentage of correct classifications of the screening system for the evaluation set for the SOM and KNN classifiers

	SGLDM	GLDS
A. SOM		
Window size: 3x3	61%	46%
Window size: 5x5	71%	67%
B. BP		
	75%	70%
C. RBF		
	65%	56%
D. KNN		
K: 9	68%	54%
K: 21	72%	65%

VI. Concluding Remarks

In conclusion, texture features provide useful information for the characterization of profusion categories in interstitial lung diseases. The screening system developed based on these feature sets gave satisfactory results. These results compare favorably with inter- and intra-reader variability. More work is currently in progress to exploit the usefulness of the approach when using a dataset containing several hundred chest radiographs.

Acknowledgements

Funding for the study was through a grant from the National Institute for Occupational Safety and Health (NIOSH), Grant 2R44 OHRRGM 03595.

References

[1] 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, Mosby-Year Book, Inc.: St Louis, MO. pp. 2116-227, 1996.

[2] *Guidelines for the use of ILO international classification of radiographs of pneumoconioses*, 1980, International Labour Office: Geneva.

[3] 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 films," *Occup. Med*., Vol. 14, pp. 537-543, 1988.

[4] Haralick R.M., Shanmugam K., Dinstein I., "Texture Features for Image Classification", *IEEE Transactions on Systems, Man., and Cybernetics,* Vol. SMC-3, pp. 610-621, Nov. 1973.

[5] Weszka J.S., Dyer C.R., Rosenfeld A., "A Comparative Study of Texture Measures for Terrain Classification", *IEEE Transactions on Systems, Man. & Cybernetics,* Vol. SMC-6, April 1976.

[6] Kohonen T., "The Self-Organizing Map", *Proceedings of the IEEE*, Vol. 78, No. 9, pp. 1464-1480, Sept. 1990.

[7] Demuth H., Beale M, *Neural Network Toolbox,* The MathWorks, Inc., Natick, Mass., USA, 1994.

[8] Tou J.T., Gonzalez R.C., *Pattern Recognition Principles*, Addison-Wesley Publishing Company, Inc., 1974.

[9] Christodoulou C:I:, Pattichis C.S., Pantziaris M., Tegos T., Nicolaides A., *Multi-feature texture analysis for the classification of carotid plaques,* CD-ROM Proceedings of the International Joint Conference on Neural Networks, IJCNN '99, Washington DC, 10-16 July, 1999.