New Cross- listed Bioengineering Course offering:
Computer Aided Diagnostics (BE420- BME520).

Instructor : **Walker H. Land, Jr. : 7-4880,**
wland@binghamton.edu
or wland6077@aol.com (home)

This is a cross listed new graduate course that will also be offered as part of the Watson School Enginet® Program.

Course Description : This is an overview course, which covers how computers can be and are currently used to aid the detection and diagnosis (CAD) of cancer as well as cancer prediction and prognosis. The fundamental goals of these two CAD processes are significantly different. We will also study a number of FDA approved computer aided diagnostic (or detection) systems in current that use several different data collection modalities. The course will have one (or more) group (or individual) projects as well as a mid term and final exam. Topical outlines for five of the main subject areas are delineated below, as well as reference to two recently published research papers in CAD. The outlines for two additional topics: FDA approved CAD systems, and time series analyses have not been completed at the time of this posting and are not included below.

Time: Fall 2008- Monday and Wednesday: 8:30am - 9:50 am. : Room J15

Call 777-5779 or 777-4880 for additional information

Objectives:

(1.) Learn the current status of CAD as it is used in medical practice today
(2.) Become familiar with how complex adaptive systems (CAS) are used in detection and diagnosis of various cancer types using different imaging modalities
(3.) Learn about FDA approved second opinion detection and diagnostic devices
(4.) Learn how CAS are also applied to cancer prediction and prognosis

Grading

Homework, Quizzes (other assignments)	**10**
Midterm	**15**
Biomedical Project(s) (group or individual)	**50**
Final Exam	**20**
Class Participation/attendance	**5**
Total	**100**

Topic 1: Computer Aided Diagnosis (CAD), and Computer Aided Detection (CAD) in General and in Breast, and Lung Cancer and Alzheimer's

- ☐ Medical Informatics Definitions
- ☐ Medical Image Processing and Analysis definitions
- ☐ What is computer aided diagnosis?

- ☐ Why should we use computer aided diagnosis?
- ☐ Computer aided detection (CAD) overview
- ☐ Medical Applications on a GRID Infractstructure Connection (MAGIC-5)
- ☐ What, exactly, is CAD and how should it be used?
- ☐ Human and computer tasks in CAD
- ☐ Where are CAD systems the most useful
- ☐ The R2 Image-Checker)
- ☐ Parts of a CAD Unit
- ☐ Software of a CAD system
- ☐ ROC review and CAD evaluation
- ☐ CAD is not generally 100% sensitive (will not get all cancers)
- ☐ CAD summary so far
- ☐ MAGIC-5 Applications of CAD to breast cancer
- ☐ Two types of CAD problems in mammography
- ☐ (1.) Microcalcifications and (2.) masses (sometimes called opacities)
- ☐ MAGIC-5 used for CAD for lung cancer
- ☐ MAGIC-5 USED In CAD for Alzheimer's

****

Paper on Image processing covering some new approaches

New Statistical Learning Theory Paradigms Adapted to Breast Cancer Diagnosis / Classification Using Image and Non-Image Clinical Data

---

Walker H. Land, Jr[1] , John J. Heine[2] , Tom Raway[1] , Alda Mizaku[1] Nataliya Kovalchuk[2] , Jack Y. Yang[3] and Mary Q. Yang [4]

[1] Dept. of Bioengineering, Binghamton Univ., Binghamton, NY, 13903-6000, USA

[2] Moffitt Cancer Center and University of South Florida Tampa, USA

[3] Harvard Medical School, Harvard University, Cambridge, Massachusetts, 02140, USA

[4] National Human Genome Research Institute, National Institute of Health, U.S. Department of Health, Bethesda, MD 20852 USA

Abstract: This paper describes three separate breast cancer research studies using both magnetic resonance mammography and screen film mammography coupled with clinical and other feature variables that addresses the following research question : how do we address the false positive biopsy artifact in diagnostic mammography, while still maintaining high sensitivities? Two new machine intelligence paradigms are used to study this problem: a Evolutionary Programming / Evolutionary Strategies Stochastic SVM hybrid (EP/ ES stochastic hybrid) and two kernelized Partial Least Squares paradigms (auto K-PLS and k-PLS). The research studies performed were: (1) performance comparison of the EP /ES SVM stochastic hybrid with the standard iterative method of training using an identical data set and statistical cross validation methods, (2) performance tradeoff and diagnostic accuracy of the auto-K-PLS, K-PLS and the EP / ES SVM hybrids using the MRM and non-image data sets, and (3)developing EP /ES SVM hybrid and K-PLS classification accuracies using the BI-RADS and clinical feature data set. These

studies showed that:(1) the new EP /ES hybrid produced comparable results to the more standard methods of iterative SVM training, but quicker, and (2) the new auto K-PLS and K-PLS systems will train and operate in essentially real time for data sets of reasonable size. These faster training and operating times are advantages in the clinical environment. Specific breast cancer classification and diagnostic research results are discussed in the paper.

Keywords: Kernel-Partial Least Squares, Evolutionary Programming / Evolutionary Strategies Derived Support Vector Machines, Machine Intelligence, computer aided diagnosis / classification

Biographical notes: **Walker H. Land, Jr.** is currently a Research Professor in the Dept. of Bioengineering as well a Principal Investigator and Director of a Computational Intelligence group there. He has over 30 years of industrial research experience and over 25 years of academic research and teaching experience. He is the author/co-author of over 200 peer reviewed research as well as several other publications.

Biographical notes: John J. Heine is currently an Associate Professor at the Moffitt Cancer Center and the University of South Florida Tampa, FL. He has over fifteen years experience in imaging physics and in the automated statistical analysis of mammograms.

Tom Raway is currently a graduate student in the Dept. Of Bioengineering at Binghamton University and a member of the Computational Intelligence research Group.

Alda Mizaku is currently a graduate student in the Dept. Of Bioengineering at Binghamton University and a member of the Computational Intelligence research Group.

Nataliya Kovalchuk is Research Associate with the Moffitt Cancer Center and University of South Florida Tampa, FL, 33620 USA.

Jack Yang is with Harvard Medical School, Harvard University, Cambridge, MA, 02140 USA

---

[4] **M. Q. Yang** is with National Human Genome Research Institute, National Institute of Health, U.S. Department of Health, Bethesda, MD 20852 USA

****

Topic 2: Survey of Automated Cancer Diagnosis (CAD) Based on Histopothological Images

- Medical Informatics Definitions
- What is Machine Learning?
- Medical Image Processing and Analysis Definitions
- What is computer aided diagnosis (CAD)
- Background
- Overview
- Examples of histopathological images
- Steps in CAD of the Images of tissue samples
- Overview of the computational steps of automated cancer diagnosis
- More on Processing steps
- Noise Reduction
- Cell segmentation
- Region based segmentation approaches
- Boundary based segmentation approaches

- Advantages / Disadvantages of the boundary and region based approaches to segmentation
- Overview of feature extraction
- Should we use the boundary or region based approaches?
- Discussion of Morphological features and computing cell parameters
- Use of morphological features
- Textural features
- Use of textural features at the cellular level
- Fractal based features
- Topological features
- Example of Voronoi diagram of a tissue sample and its Delaunay triangulation
- Example of constructing a cell graph of an image
- Intensity based features and Intensity based features with (color) histograms
- Feature measurement and selection
- Classification for detection and diagnosis
- Machine learning algorithm classifiers that have been used in detection / diagnosis
- Evaluation of classifiers for detection / diagnosis
- Some complex adaptive systems (CAS) classification approaches that have been used
- Types of features that have been used in classifying different cancer types

****

Related paper on the same topic:

A kernelised fuzzy-Support Vector Machine CAD system for the diagnosis of lung cancer from tissue images

Walker H. Land, Jr.*

Department of Bioengineering, Binghamton University,
Binghamton, NY, 13903-6000, USA
E-mail: wland@binghamton.edu
*Corresponding author

Daniel W. McKee

Department of Mathematics
and Computer and Information Sciences,
Mansfield University, Mansfield, PA, 16933, USA
E-mail: dmckee@mansfield.edu

Tatyana Zhukov, Dansheng Song
and Wei Qian

Department of Interdisciplinary Oncology,
Colleges of Medicine and H. Lee Moffitt Cancer Center
and Research Institute,
University of South Florida, Tampa, FL, 33612, USA
E-mail: Zhukov@moffitt.usf.edu
E-mail: dsong2@mail.usf.edu
E-mail: wqian@health.usf.edu

Abstract: This research describes a non-interactive process that applies several forms of computational intelligence to classifying biopsy lung tissue samples. Three types of lung cancer evaluated (squamous cell carcinoma, adenocarcinoma, and bronchioalveolar carcinoma) together account for 65–70% of diagnoses. Accuracy achieved supports hypothesis that an accurate predictive model is generated from training images, and performance achieved is an accurate baseline for the process's potential scaling to larger datasets. Feature vector performance is good or better than Thiran and Macq's in every case. Except bronchioalveolar carcinomas, each individual cancer classification task experienced improvement, with two groupings showing nearly 20% classification accuracy.

Keywords: Computer-Aided Diagnosis; CAD; lung cancer; segmentation; feature selection; classification; microscopy images; kernel methods; Support Vector Machine; SVM.

Copyright © 200x Inderscience Enterprises Ltd.

****

Topic 3: Microarray Background and most recent application of microarray genomic technology to the study of breast cancer

- ☐ Overview
- ☐ DNA technical background
- ☐ Why microarrays are important
- ☐ What is a microarray?
- ☐ Designing microarray experiments
- ☐ Example of an illuminated ,enlarged microarray
- ☐ Microarray processing graphic
- ☐ Summary of experimental approach
- ☐ Comparative Genomic Hybridization (CGH)
- ☐ Changes in gene expression levels
- ☐ DNA repair genes

- ☐ DNA mutations
- ☐ Some microarray applications areas
- ☐ Application of microarrays to treatment / management of breast cancer
- ☐ Focus is application genomic technology to studying breast cancer
- ☐ Summary
- ☐ Overview
- ☐ Background
- ☐ Molecular classification of breast cancer
- ☐ Prognostic Implications
- ☐ Genetic susceptibility
- ☐ Predictive implications
- ☐ Conclusions

****

Topic 4: CAD and machine learning as applied to cancer Prediction and Prognosis

- **Overview**
- **Survey of CAD methods applied to cancer prediction approach**
- **Cruz and Wishart Case study 1: Cancer Risk or "Susceptibility" Prediction**
- **Cruz and Wishart Case study 2: Prediction of cancer survivability**
- **Cruz and Wishart Case study 3-Prediction of cancer recurrence**
- **Lessons Learned**
- **Conclusions**
- **References**

Topic 5: Introduction and overview of genomics. Application of genomic training data to CAD of cancer prediction and prognosis

- **Goal of molecular biology**
- **What is a genome?**
- **Physical structure of the human genome**
- **Two complementary DNA polypeptides**
- **Mitochondrial DNA**
- **Why study mitochondrial DNA?**
- **Ribonucleic acid (RNA) revisited**
- **Genes and protein synthesis (transcription, splicing and translation)**
- **Eight referenced cancer research studies that used genomic training data for several different cancers with different measured clinical outcomes**

Important Points you need to know about this class (Please read this information carefully):

- ***Accommodations***

  If you are a student with a disability and wish to request accommodations, please notify me by the second week of class. You are also encouraged to contact the Office of Services for Students with Disabilities (SSD) at 777-2868. Their office is at LH-B51. The SDD office makes formal recommendations regarding necessary and

appropriate accommodations based on your specifically diagnosed disability. Information regarding your disability will be treated in a confidential manner.

- *The class will be taught* ***using the instructor lecture notes*** *Consequently, it is imperative that you attend class to understand the full meaning of the material covered as well as to understand how to apply the material to the solution of practical problems. Attendance will be taken.*

- *Projects will require you to do literature research but are designed to eliminate any required knowledge of the mathematical complexities associated with the design and implementation of complex adaptive systems . The only mathematical processing required will be to develop ROC curves, and a lecture will be devoted to explain the details of that process, which will include several examples.*

- ***The projects will require a professional report and presentation****. A detailed outline (for both the report and presentation) will be provided (unless you choose your own individual or research group project, which is subject to the approval of the instructor.) Class members will form their own groups and select a group leader from that group. (Please note that this arrangement is for convenience of communication with the group and does not restrict any student form talking to the instructor at any time.)*

- *Homework and other types of assignments (other than in class assignments) will be usually be due one week after assignment. No late work will be accepted, .other than that accompanied by a valid excuse: Example: physician provided excuse for illness. Furthermore, Homework will be graded on my ability to understand that you are learning from the material, not on whether you show a "right" or "wrong" answer. Therefore, clarity in presentation of your solutions is essential. Homework should be submitted in hardcopy form or* ***by other means as specified****.*

- ***Incidents of academic dishonesty will be fully investigated and processed in accordance with university regulations. As teamwork is a requirement for the group projects, you should/must always give full credit where credit is due as well as complete your assigned share of the project work.***