GRANT NUMBER: DAMD17-94-J-4371

TITLE: Computer Aided Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Carey E. Floyd, Jr., Ph.D.

CONTRACTING ORGANIZATION: Duke University Medical Center
Durham, NC 27710

REPORT DATE: February 1996

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, MD 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
Computer Aided Breast Cancer Diagnosis

Carey E. Floyd, Jr., Ph.D.

Duke University Medical Center
Durham, NC 27710

Commander
U.S. Army Medical Research and Materiel Command,
Fort Detrick, Frederick, MD 21702-5012

Approved for public release; distribution unlimited

The long range goal of this project is to improve the accuracy and consistency of breast cancer diagnosis by developing a Computer Aided Diagnosis (CAD) system for early prediction of breast cancer from the patients' mammographic findings and medical history.

In this progress period we have developed an ANN to predict biopsy outcome from mammographic and history findings. In the first year of the grant we have 1) published the preliminary data presented in the original proposal, 2) acquired 260 cases using the standardized BI-RADS reporting system and published this study, 3) conducted a small prospective study, 4) examined the inter- and intra-observer variability of the reporting lexicon, and 5) investigated reducing the number of active input features to the network and published this work. A related study was conducted to examine the sensitivity of the system to the techniques used for sampling the data. All of this work has been specifically directed toward the first specific aim of the proposal.

Computer Aided Diagnosis, Mammography, Breast Cancer

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102
The Report Documentation Page (RDP) is used in manuscript and manuscript review. It is important that this information be consistent with the rest of the report, particularly the abstract and title page. Instructions for filling in each block of the form follow. It is important to complete the block to meet optical scanning requirements.

Block 1. **Agency**: Use Only (Leave Blank).

Block 2. **Report Date**: Full publication date including day, month, and year, if available (e.g., 1 Jan 88). Must cite at least one year.

Block 3. **Type of Report and Dates Covered**: State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g., 10 Jun 87 - 30 Jun 88).

Block 4. **Title and Subtitle**: A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. **Funding Numbers**: To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

- C = Contract
- G = Grant
- PE = Program
- PR = Project
- VA = Task
- WU = Work Unit
- Element
- Accession No.

Block 6. **Author(s)**: Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. **Performing Organization Name(s) and Address(es)**: Self-explanatory.

Block 8. **Performing Organization Report Number**: Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. **Sponsoring/Monitoring Agency Name(s) and Address(es)**: Self-explanatory.


Block 11. **Supplementary Notes**: Enter information not included elsewhere such as:

- Prepared in cooperation with...
- Trans. of...
- To be published...
- When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12. **Abstract**: Limited to a brief (100 words) 250-word factual summary of the most significant information contained in the report.

Block 13. **Subject Terms**: Key words or phrases identifying major subjects in the report.

Block 14. **Number of Pages**: Enter the total number of pages.

Block 15. **Price Code**: Enter appropriate price code (NLM code).

Block 16. **Security Classification**: Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 17. **Limitation of Abstract**: This block must be completed to assign a limitation to the abstract. Enter either III (unclassified) or SFI (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unclassified.

---

*Standard Form 250 Back (Rev. 2-80)*
FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Where copyrighted material is quoted, permission has been obtained to use such material.

Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

For the protection of human subjects, the investigator(s) have adhered to policies of applicable Federal Law 32 CFR 219 and 45 CFR 46.

In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

[Signature]
Principal Investigator's Signature

[Date]
Table of Contents:

Front cover  
1  
SF298  
2  
Foreword  
3  
Table of Contents  
4  
Introduction  
5  
Body  
6
INTRODUCTION

The long range goal of this project is to improve the accuracy and consistency of breast cancer diagnosis by developing a Computer Aided Diagnosis (CAD) system for early prediction of breast cancer from the patients’ mammographic findings and medical history.

While mammography is a sensitive test for early diagnosis of breast cancer, 70% of the cases which are sent to biopsy are benign. We will develop a CAD system based on Artificial Neural Networks (ANNs) to predict the malignancy of breast findings (as verified by the outcome of biopsy) based upon radiologists' findings from mammograms. The strength of ANNs for this problem is their ability to learn complex relationships from examples of the data, then to generalize and accurately classify examples which the network has not seen before. This system will learn to predict malignancy by examining a large set of radiographic findings which are paired with biopsy results. The database for this learning will be representative of the patient population. Specifically we will, 1) Develop an ANN to predict biopsy outcome from mammographic and history findings. 2) Evaluate the improvement in radiologists' diagnostic performance when the computer diagnostic aid is provided. This implementation of an accurate CAD system will improve sensitivity, specificity, and consistency of breast cancer diagnosis and will provide a significant improvement in long term outcome for breast cancer patients.
BODY

In the first year of the grant we have published 3 peer-reviewed manuscripts[1-3] with two more accepted for publication[4][5]. Another manuscript has been submitted for peer review[6Tourassi]. There have been four presentations with published proceedings at professional meetings [7][8][9][10]. In addition, we have applied for and received funding for supplemental funding to extend work on this grant[11]. Specifically, we have 1) published the preliminary data presented in the original proposal[1], 2) acquired 260 cases using the standardized BI-RADS reporting system and published this study[2], 3) conducted a small prospective study (accepted for publication[4]), 4) examined the inter- and intra-observer variability of the reporting lexicon (accepted for publication[5]), and 5) investigated reducing the number of active input features to the network and published this work[3][7][9]. A related study was conducted to examine the sensitivity of the system to the techniques used for sampling the data[6][8]. A genetic algorithm was developed as an alternative approach to the ANN [10]. All of this work has been specifically directed toward the first specific aim of the proposal.

In summary:

Year 1:

Peer-reviewed manuscripts published or in press: 5
Peer-reviewed manuscripts submitted: 1
Published Conference Proceedings: 4
International Meeting presentations: 7
Supplemental grants received: 1

Peer-reviewed manuscripts published or in press:


**Peer-reviewed manuscripts submitted:**


**Published Conference Proceedings:**


Meeting presentations:


