# DATA MINING APPROACH TO IMAGE INFORMATION: FEATURE SELECTION IN MAMMOGRAM MICROCALCIFICATIONS


## 吳智誠、傅家啟,李三剛

E-mail: 9018609@mail.dyu.edu.tw



**ABSTRACT**

THE MICROCALCIFICATIONS IN X-RAY MAMMOGRAPHY ARE ONE OF THE EARLY MAJOR INDEX OF BREAST CANCER. THE COMPUTER AIDED DIAGNOSIS (CAD) CAN BE USED TO DETECT MICROCALCIFICATIONS AS A SECOND OPINION STRATEGY. THE PUBLISHED LITERATURE RELATED TO CAD DEVELOPMENT STILL FOLLOWED THE CONVENTIONAL ENGINEERING TECHNIQUES. HOWEVER, THE PERFORMANCE OF A CAD FOR MAMMOGRAM MICROCALCIFICATIONS IS HEAVY INFLUENCED BY 1. THE INSPECTION MODES, 2. THE CHOICE OF SUB-SYSTEM (E.G. IMAGE ENHANCEMENT), AND 3. THE METHODS OF FEATURE EXTRACTION AND THE COMBINATIONS OF FEATURE VECTOR. THE CURRENT PERFORMANCE ANALYSIS CAN NOT EFFECTIVELY ANSWER THE FOLLOWING TWO QUESTIONS: 1. HOW TO DECIDE THE BEST COMBINATION OF SUB-SYSTEMS, AND 2. HOW TO SELECT THE BEST COMBINATION OF FEATURES. IN THIS THESIS, THE DATA MINING TECHNIQUES ARE APPLIED TO THE MEMORY-BASED REASONING (MBR) INSPECTION SYSTEM TO OVERCOME THE AFORE-MENTIONED PROBLEMS. THE FIRST PHASE IS TO DECIDE IMAGE ENHANCEMENT OPERATING PARAMETERS OF THE SCANNING INSPECTION MODE, AND THE SECOND PHASE IS TO APPLY DATA MINING BASED FEATURE SELECTION APPROACH TO DELETE THE REDUNDANT FEATURES TO IMPROVE THE EFFICIENCY OF THE SYSTEM. EXPERIMENTAL RESULTS SHOWED THAT DATA MINING APPROACH PROVIDES SATISFACTORY PERFORMANCE BY REDUCING REDUNDANT FEATURES.



Keywords : DATA MINING, FEATURE SELECTION, COMPUTER AIDED DIAGNOSIS, MEMORY-BASED REASONING, MICROCALCIFICATIONS, X-RAY MAMMOGRAPHY


**Table of Contents**


第一章 緒論--P1 1.1研究目的與動機--P1 1.2研究範圍--P1 1.3研究方法--P2 第二章 文獻探討--P3 2.1微鈣化檢測--P3 2.1.1整張數位化X光乳房影像微鈣化檢測分類模式--P5 2.1.2影像區塊掃瞄微鈣化檢測模式--P8 2.2特徵選擇相關議題探討--P9 2.2.1特徵選擇的定義--P9 2.2.2特徵選擇的步驟--P9 2.2.3 GENERAL REGRESSION NEURAL NETWORK--P10 2.2.3.1 GRNN--P11 2.2.3.2 GRNN FEATURE SELECTION--P12 2.3績效衡量-ROC曲線--P13 第三章 研究方法--P16 3.1資料探勘--P16 3.2前處理-小波變換影像強化--P18 3.3資料轉換-空間轉換特徵萃取--P22 3.3.1 GRAY-TONE SPATIAL-DEPENDENCE MATRICES模式--P23 3.3.2 GRAY LEVEL RUN LENGTHS METHOD模式--P24 3.3.3 GRAY LEVEL DIFFERENCE METHOD模式--P24 3.4特徵選擇與分類器--P25 3.4.1記憶基礎理解--P25 3.4.2特徵選擇與分類器--P26 第四章 實驗結果--P28 4.1實驗設置及流程--P28 4.1.1訓練資料集擷取--P30 4.1.2特徵萃取--P30 4.1.3影像強化評量--P32 4.1.4特徵選擇與績效評量--P33 4.2 實驗結果及分析--P34 4.2.1影像強化評量結果--P34 4.2.2特徵選擇結果--P37 4.2.3績效衡量--P38 4.2.4實驗討論--P40 第五章 結論與未來研究發展--P43 5.1 結論--P43 5.1.1強化模式衡量探討--P43 5.1.2特徵選擇衡量探討--P43 5.2 未來研究發展--P44 參考文獻--P45 附錄A--P47 附錄B--P51 附錄C--P52 附錄D--P53 附錄E--P60 附錄F--P67


**REFERENCES**


1.彭文正,"資料採礦-顧客關係管理暨電子行銷之應用" ,數博網資訊股份有限公司,台北, 2001 2.羅見順, 國立成功大學電機工程所博士班論文"乳房X光攝影之微鈣化群電腦輔助診斷系統", 2000 3.ARIAN, R. E, AND PETER M.TH. PATTYNAMA, "RECEIVER OPERATING CHARACTERISTIC (ROC) ANALYS IS: BASIC PRINCIPLES AND APPLICATION IN RADIOLOGY", EUROPEAN JOURNAL OF RADIOLOGY 27, P88-94, 1998 4.CONNERS, R. W., HARLOW, C. A., "A THEORETICAL COMPARISON OF TEXTURE ALGORITHMS", IEEE TRANS. PATTERN ANAL. MACH. INTELL. PAMI-2, NO. 3 , P204-222, 1980 5.DASH M., LIU H., "FEATURE SELECTION FOR CLASSIFICATION", INTELLIGENT DATA ANALYSIS (1), P131-156, 1998 6.GALLOWAY, M.M., "TEXTURE ANALYSIS USING GRAY RUN LENGTHS", COMPUTER GRAPHICS AND IMAGE PROCESSING , P172-179,1975 7.JOACHIM DENGLER, SABINE BEHRENS, AND JOHANN FRIEDRICH DESAGA, "SEGMENTATION OF MICROCAL -CIFICATIONS IN

MAMMOGRAM", IEEE TRANSACTION ON MEDICAL IMAGING, VOL.12, NO.4, DECEMBER 1993 8.KIM, J. K., PARK, H.W. "STATISTICAL TEXTURAL FEATURES FOR DETECTION OF MICROCALCIFICATI -ONS IN DIGITIZED MAMMOGRAMS", IEEE TRANSACTION ON MEDICAL IMAGING, VOL.18, NO.3, MARCH 1999. 9.KARSSEMEIJER, N., "ADAPTIVE NOISE EQUALIZATION AND IMAGE ANALYSIS IN MAMOGRAPHY",PROCEE -DINGS INFORMATION PROCESSING IN MEDICAL IMAGING (IPMI) , FLAGSTAFF , JUNE 1993 10.LIU, H., HIROSHI MOTODA, "FEATURE SELECTION FOR KNOWLEDGE DISCOVERY AND DATA MINING" , KULWER ACADEMIC PUBLISHERS 11.NISHKAWA, R.M. , GIGER, M.L. ,DOI K.,VYBORNY C.J.,SCHMIDT R.A., "COMPUTER-AIDED DETECTI -ON OF CLUSTERED MICROCALCIFICATION ON DIGITAL MAMMOGRAMS" ,MEDICAL PHYSICS AND IMAGING ,1995 12.ROBERT, M.H., SHANMUGAM, K. ,AND ITS'HAK DINSTEIN, "TEXTURE FEATURES FOR IMAGE CLASSIF -ICATION" ,IEEE TRANS.ON SYSTEM, MAN, AND CYBERNETICS, VOL. SMC-3, NO.6 ,NOVEMBER 1973 13.SPECHT, D. F., "A GENERAL REGRESSION NEURAL NETWORK", IEEE TRANSACTION ON NEURAL NETWOR -K. VOL. 2., NO. 6., NOVEMBER 1991 14.STRICKLAND,R.N. ,HAHN,H.I. ,"WAVELET TRANSFORMS FOR DETECTION MICROCALCIFICATIONS IN MAMMOGRAMS", IEEE TRANS. ON MEDICAL IMAGING, VOL.15,NO.2,APRIL 1996. 15.YU, S.Y. AND GUAN L., "A CAD SYSTEM FOR THE AUTOMATIC DETECTION OF CLUSTERED MICROCALC -IFICATION IN DIGITIZED MAMMOGRAM FILMS ", IEEE TRANS. OF MEDICAL IMAGING , VOL. 19, FEBRUARY 2000 16.ZHENG B.Y., QIAN WEI, AND CLARKE L. P., "DIGITAL MAMMOGRAPHY: MIXED FEATURE NEURAL NETWORK WITH SPECTRAL ENTROPY DECISION FOR DETECTION OF MICROCALCIFICATIONS" , IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 15,NO. 5, OCTOBER 1996