UNIVERSITI PUTRA MALAYSIA

EFFECTIVENESS OF TREE BASED PIPELINE OPTIMIZATION TOOLS AND GRID SEARCH METHOD IN BREAST CANCER PREDICTION

SITI FAIRUZ BINTI MAT RADZI

FS 2022 4

EFFECTIVENESS OF TREE BASED PIPELINE OPTIMIZATION TOOLS AND GRID SEARCH METHOD IN BREAST CANCER PREDICTION

By

SITI FAIRUZ BINTI MAT RADZI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

October 2021

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment ofthe requirement for the degree of Master of Science

EFFECTIVENESS OF TREE BASED PIPELINE OPTIMIZATION TOOLS AND GRID SEARCH METHOD IN BREAST CANCER PREDICTION

By

SITI FAIRUZ BINTI MAT RADZI

October 2021

Chair : Dr. Muhammad Khalis Bin Abdul Karim, PhD
Faculty : Science

Breast cancer has been known as the most prevalent and common cause of death among Malaysian woman especially over the age of 40. Breast cancer can usually be identified as either benign or malignant with invasive biopsy procedure. The treatment protocol is allocated based on the whether the mass is benign or malignant. Fortunately, breast cancer like many other cancer types are curable and patient survival can be improved, subject to early diagnosis. Radiograph images lies numbers of features that useful for computer aided diagnosis.In this thesis, the work is divided into two main phases; 1) evaluating the reproducibility of radiomics features derived from manual delineation and semiautomatic segmentation after two different contrast enhancement techniques on masses in two-dimensional (2D) mammography images and 2) to implement the Automated Machine Learning (AutoML) in classifying types of mass in mammogram images. With introduction of ML techniques, breast cancer can be diagnosed in early stage without any invasive and risky procedure. The methodology presented in this research consist of several stages including, image acquisition, image segmentation, feature extraction/selection and, classification using AutoML. The first phase determines the reproducibility between Contrast Limited Adaptive Histogram Equalization (CLAHE) and Adaptive Histogram Equalization (AHE) techniques. The semiautomatic segmentation techniques used in the first phase is Active Contour Method (ACM) with 100 iterations. Three types of radiomics features were extracted including first order, second order and shape

features. 37 features were extracted from each tumor in three different techniques mentioned: 9 of these were shape-based features, while 28 were texture-based features.Notably the CLAHE group (ICC = 0.890 ± 0.554, $p < 0.05$) had the highest reproducibility compared to the features extracted from the AHE group (ICC = 0.850 ± 0.933, $p < 0.05$) and manual delineation (ICC = 0.673 ± 0.807, $p > 0.05$). Therefore, the segmentation techniques used in the second phase are based on CLAHE and ACM method. The Principal Component Analysis (PCA) Random Forest (RF) classification has proved to be the most reliable pipelines with the lowest complexity in this researchwith 92% of accuracy, 83% of precision, 100% of sensitivity, 94% of ROC.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

KEBERKESANAN PENGOPTIMUMAN ALAT TALIAN PAPIP BERASASKAN POKOK DAN KAEDAH PENCARIAN GRID DALAM RAMALAN KANSER PAYUDARA

Oleh

SITI FAIRUZ BINTI MAT RADZI

Oktober 2021

Pengerusi : Dr. Muhammad Khalis Bin Abdul Karim, PhD
Fakulti : Fakulti Sains

Kanser payudara telah dikenali sebagai penyebab kematian yang paling lazim dan biasa di kalangan wanita Malaysia terutamanya yang berusia lebih dari 40 tahun. Kanser payudara biasanya dapat dikenal pasti sebagai benigna atau malignan dengan prosedur biopsi yang invasif. Protokol rawatan diperuntukkan berdasarkan sama ada ketulan itu merupakan benigna atau malignan. Namun begitu, kanser payudara seperti jenis barah lain dapat disembuhkan dan kelangsungan hidup pesakit dapat ditingkatkan, bergantung pada diagnosis awal. Imej radiograf mengandungi sebilangan besar ciri yang berguna untuk diagnosis berbantukan komputer. Dalam tesis ini, kajian terbahagi kepada dua fasa utama; 1) menilai kebolehulangan ciri radiomik yang berasal dari persempadanan manual dan segmentasi semiautomatik setelah dua teknik peningkatan kontras yang berbeza pada ketulan dalam imej mamografi dua dimensi (2D) dan 2) untuk menerapkan Pembelajaran Mesin Automatik (AutoML) dalam mengklasifikasikan jenis ketulan dalam imej mamogram.Dengan pengenalan teknik ML, barah payudara dapat didiagnosis pada peringkat awal tanpa prosedur invasif dan berisiko. Metodologi yang dikemukakan dalam penyelidikan ini terdiri dari beberapa tahap termasuk, pemerolehan gambar, segmentasi gambar, pengekstrakan / pemilihan tapisan dan, klasifikasi menggunakan AutoML. Fasa pertama menentukan kebolehulangan antara teknik Contrast Limited Adaptive Histogram

Equalization (CLAHE) dan Adaptive Histogram Equalization (AHE). Teknik segmentasi semiautomatik yang digunakan pada fasa pertama adalah Kaedah Kontur Aktif (ACM) dengan 100 lelaran. Tiga jenis ciri radiomik diekstraksi termasuk urutan pertama, susunan kedua dan ciri bentuk. 37 ciri diekstrak dari setiap tumor dalam tiga teknik berbeza yang disebutkan: 9 daripadanya adalah ciri berdasarkan bentuk, sementara 28 daripadanya adalah ciri berasaskan tekstur. Terutama kumpulan CLAHE (ICC = 0.890 ± 0.554, $p < 0.05$) mempunyai kebolehulangan tertinggi berbanding dengan ciri yang diekstrak dari kumpulan AHE (ICC = 0.850 ± 0.933, $p < 0.05$) dan penerapan teknik manual (ICC = 0.673 ± 0.807, $p > 0.05$). Oleh itu, teknik segmentasi yang digunakan pada fasa kedua adalah berdasarkan kaedah CLAHE dan ACM. Klasifikasi Analisis Komponen Utama (PCA)Ekstra Pokok (ET) telah terbukti sebagai saluran paip yang paling dipercayai dengan kerumitan terendah dalam penyelidikan ini dengan 92% ketepatan, 83% ketepatan,100% kepekaan , 94% ROC.

ACKNOWLEDGEMENTS

All praises to Allah, the Most Gracious, the Most Merciful and thousands upon Hislast messenger, Muhammad S.A.W. I am grateful to be able to complete this research within given period of time.
I would like to express my gratitution to my supervisor, Dr. Muhammad Khalis Bin Abdul Karim for his never ending guidance and his continuous support throughout my research journey. I would like to thank my co-supervisors, Dr. Mohd Amiruddin Bin Abd Rahman and Prof. Dr. Iqbal Saripan, for their assistance and guidance. They have helped me tremendously especially in improving my work.
Last, but not least, I am greatly indebted to my family and friends, especially to my parent, Azimah Binti Abdul Aziz. With her unconditional love and constructive motivation, the hardship throughout my research were lighten.
The research for this thesis was financially funded and supported by the Ministry of Higher Education Grant under Grant FRGS/5540112. Finally, special thanks to all parties and Universiti Putra Malaysia for their facilities and funding for this research.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Muhammad Khalis Bin Abdul Karim, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Mohd Amiruddin Bin Abd Rahman, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

M. Iqbal Bin Saripan, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

_____
ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 10 February 2022

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of Chairman of Supervisory Committee: Muhammad Khalis Bin Abdul Karim

Signature: _____
Name of Member of Supervisory Committee: Mohd Amiruddin Bin Abd Rahman

Signature: _____
Name of Member of Supervisory Committee: M. Iqbal Saripan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MNCR Registry	Malaysia National Cancer
CR	Cumulative Risk
ASR	Age Standardized Rate
CumR	Cumulative Rate
ML	Machine Learning
AutoML Learning	Automated Machine
GS	Grid Search
TPOT	Tree-Based Pipeline Optimization Tools
DDSM	Digital Database for Screening Mammography
CBIS-DDSM	Curated Breast Imaging Subset of DDSM
TCIA	The Cancer Imaging Archive
MLO	Medio-Lateral Oblique
CLAHE	Contrast Limited Adaptive Histogram Equalization
AHE	Adaptive Histogram Equalization
BBHE	Brightness Bi-Histogram Equalization
DSIHE	Dualistic Subimage Histogram Equalization
RMSHE	Recursive Mean Separate Histogram Equalization
MMBEBHE	Minimum Mean Brightness Error Bi-Histogram Equalization
RWSHE	Recursive Separated and Weighted Histogram Equalization
GLSZM	Gray Level Size Zone Matrix
GLRLM	Gray Level Run Length Matrix

NGTDM	Neighbouring Gray Tone Difference Matrix
GLDM	Gray Level Dependence Matrix
ATM	Auto-Tuned Models
ROI	Region of Interest
SVM	Support Vector Machine
NB	Naive Bayes
MLP-ANN	Multi-Layer Perceptron-Artificial Neural Network
CC	Cranioaudal
FFDM	Full-Field Digital Mammogram
DBT	Digital Breast Tomosynthesis
HDTV	High Definition Television
HE	Histogram Equalization
ANCE	Adaptive Neighbourhood Contrast Enhancement
ACM	Active Contour Models
GAC	Geometric Active Contour model
ACWE	Active Contour Without Edges
GLCM	Gray Level Co-occurrence Matrix
RFE	Recursive Feature Elimination
PSO	Particle Swarm Optimization
DT	Decision Tree
LR	Logistic Regression
FSS	Feature Set Selector
DICOM	Digital Imaging and Communications in Medicine

BI-RADS	Breast Imaging-Reporting and Data System
SPSS	Statistical Package for the Social Sciences
MATLAB	Matrix Laboratory
CL	Clip Limit
ANN	Artificial Neural Network
AUC	Area Under the Curve
ICC	Intra and Inter-class Correlation Coefficient
ANOVA	Analysis of Variance
MS_R	mean square for rows
MS_E	mean square error
MS_C	mean square for columns
TP	True Positive
FN	False Negative
TN	True Negative
FP	False Positive
PPV	positive predictive value
ROC	Receiver Operating Characteristics
CAD	Computer Aided Diagnosis
CT	Computed Tomography
GP	Genetic Programming
PCA	Principal Component Analysis
RF	Random Forest
SE	Stacking Estimator

SS	Standard Scaler
CASH	Combined Algorithm Selection and Hyperparameter Optimization

CHAPTER 1

INTRODUCTION

1.1 Research Background

Breast cancer has been acknowledged as the most prevalent and common cause of death among Malaysian woman over the age of 40 (Azizah *et al*., 2019). Several studies emphasize the need and urgency for early detection in reducing breast cancer morbidity and mortality (Eddy et al., 1988; Seely et al., 2018; Lambin et al., 2012).
Medical imaging techniques, such as mammography, play an important role in non-invasively assessing breast tissues for detection, diagnostic, staging, and management purposes (Seely *et al*., 2018). In an attempt to improve the mortality rate among the population, a mammography screening program is proven to be the most cost-effective program for providing useful details about the presence of abnormal mass or tumor (Seely *et al*., 2018).

According to Malaysia National Cancer Registry (MNCR) Report, release a report about breast cancer every 5 years. For 2007 to 2011, 18206 cases of female breast cancer were recorded. However the number of cases increased to 21634 from 2012 until 2016 compared to previous report. The next edition is yet to be published which covers the report from 2017 to 2021. Breast cancer accounted for 34.1% of all cancer among females in Malaysia .Over 47% cases were detected at later stage; stage 3 and stage 4. In 2007-2011 report, the percentage of cases detected at later stage was higher compared to 43.2% (Azizah *et al*., 2019). This is due to the density of breast tissue in younger women, which enable mammogram to detect the lesion accurately

Study shows that higher breast density on mammography is strongly associated withan increased risk of breast cancer, this occurs especially in younger women Vachon et al., 2007; Pinsky et al., 2010; Boyd et al., 2007).

Table 1.1 below summarized the incidence of breast cancer by year among female in Malaysia. Cumulative Risk (CR), Age Standardized Rate (ASR) and Cumulative Rate (CumR) were included in the table for each residents.

Table 1.1: Female Breast: Cancer incidence summary by year, Malaysia(Azizah *et al*., 2019)

All residents	No. of cases	Age Standardized Rate (ASR)	Cumulative Rate (CumR)
2007-2011	18206	31.1	3.4
2012-2016	21634	34.1	3.7
2012	4266	33.5	3.6
2013	4076	31.0	3.4
2014	4150	31.7	3.5
2015	4518	33.4	3.6
2016	4624	34.4	3.8

Figure 1.1 illustrates the graph of the comparison of age-specific incidence rate by year from birth to over 75 years old in Malaysia.

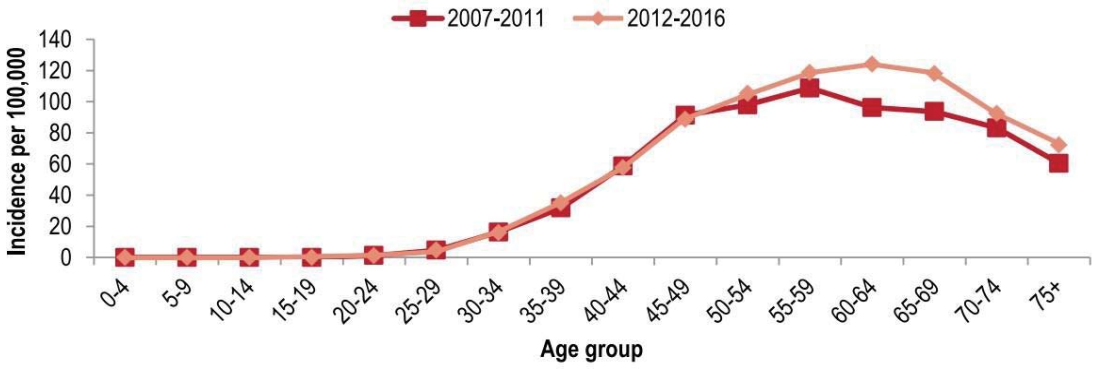

Figure 1.1: Female Breast: Comparison of age- specific incidence rate by year in Malaysia (Azizah *et al*., 2019)

The age-specific incidence rate by year for 2012-2016 recorded an increment compared to 2007-2011. The increment involving false-positive rates is a huge threat in cancer diagnosis (Nelson et al., 2016) that usually confirmed through various ways such as biopsy techniques. Only 4% to 5% of positive mammograms recalled for further evaluation ultimately lead to a cancer diagnosis (Lehman *et al*., 2017). Hence, diagnosing early stage of breast cancer accurately is very crucial sinceearly treatment can be given to the patient.

Radiomics is a term increasingly used in oncological radiotherapy in order to better appreciate the region or volume of interest (target volumes and critical organs) but also to assess the somatic or constitutional biological component. Radiomics a high-performance qualitative and quantitative analysis, consisting in the high-speed extraction of digital medical imaging data to obtain predictive and prognostic information concerning patients treated for a cancerous pathology. Via the principle of image recognition and Machine Learning (ML), computerized systems provide the opportunity to acquire knowledge about the issue in a manner that is impossible for a human being to obtain. In other words, this information could sometimes be indistinguishable by human vision (Fabijańska *et al.*, 2009).

Medical images are a powerful tool to diagnose and analyze many diseases such as breast, chest, abdominal illnesses, and blood disorder. The digital format of the medical images offers incentive for further analysis that may help to improve the accuracy of breast cancer diagnosis and hence, help to optimize the management of patient. Major contribution of image processing and machine learning techniques in the medicine field are through the digitized medical images where they can beexplored without human limitation.

1.2 Problem Statement

The main markers for breast cancer in mammograms are masses and microcalcification. Mass is defined as a space-occupying lesion, visible in two different projections, characteristic by its shape and contour (Berment *et al*., 2014) while microcalcification is defined as deposits of calcium in the breast tissue and appear as small bright spots on mammograms (Azam *et al*., 2021). Interpretation of these anomalies is a challenge due to their low mortalities (Birdwell *et al*., 2009). The enormous volume of mammogram generated by widespread screening often overwhelmed radiologists (Rangayyan *et al*., 2007), and even experienced radiologists have significant inter-observer and intra-observer variability in their mammograms interpretation (Skaane *et al*., 1997). It is even harder to identify mammographic masses than microcalcification, because masses differ greatly in shape, margin, size and typically have obscure boundaries. (Azam *et al*., 2021). Subsequently, radiologists miss a large portion of retrospectively observable masses (Birdwell *et al*., 2001), and biopsies are often performed on normal tissues and benign lesions (Hubbard *et al*., 2011). It is commonly agreed that by double reading, the sensitivity can be improved without increasing recall rates (Blanks et al., 1998), but it could be expensive due to increase in manpower. Therefore, ML method used have been developed for breast cancer detection and classification. This method is to facilitate interpretation and analysis, the preprocessing of mammography films helps improve the visibility of peripheral areas and intensity distribution.

1.3 Research Objectives

This study embarks on the following objectives:

1. To quantify and extract the radiomics feature from the segmentation of the masses in breast of 2D breast mammogram.
2. To compare the performance of semiautomatic and manual segmentation techniques of masses.
3. To evaluate Automated Machine Learning (AutoML) and Grid Search (GS)algorithm for selection of optimum features extracted from the images in orderto correctly classify mass into benign or malignant.
4. To determine the performance of Automated Machine Learning (AutoML) and Grid Search algorithm based on performance metrics.

1.4 Significance of Study

The findings of this study can contribute to a real-life scenario case. A radiologists observe mammogram to detect the mass in breast. In many cases, even experienced radiologists, might have difficulties to precisely determine the region-of-interest (ROI) in a mass; benign or malignant and often have different opinion on the location of the mass. Moreover, the radiologists still misinterpret between 10% and 30% of cancer (Ekpo *et al*., 2018)

The goal of this work is to utilize image processing and ML techniques in order to increase the accuracy of diagnosing breast cancer. Therefore, to achieve the intended goal, the research is carried out in four main stages, namely, 1. image acquisition, 2. image segmentation, 3. feature extraction and selection, and finally 4. classification.

These form the four main modules of a typical architecture of a Computer-Aided Diagnosis (CAD) system.

This research makes several key contributions as follows:

- Improvement of intra and inter-observer variability by observing the stability of extracted features using semi-automatic segmentation compared to manual segmentation.
- Comparative study with two optimization methods, Tree-Based Pipelines Optimization Tools (TPOT) algorithm and Grid Search (GS) algorithm in achieving high accuracy, sensitivity and specificity.
- The proposed approach achieves remarkable results of 92% accuracy in classifying masses in breast cancer with limited effort and time.

1.5 Scope of Thesis

The scope of study involves evaluating and extracting radiomics features as well as comparing accuracy in detecting masses in breast cancer using two optimization techniques which are Tree Based Pipeline Optimization Tools (TPOT) and GS Algorithm in order to achieve higher accuracy in less complex pipeline. To achieve this, the study was divided into two parts;

1. Part I: The updated version of Digital Database for Screening Mammography (DDSM), Curated Breast Imaging Subset of DDSM (CBIS-DDSM) data from The Cancer Imaging Archive (TCIA) open source was adopted in this research.
   30 benign mammogram images medioliteral Oblique (MLO) views were enhanced by using two techniques of contrast enhancement, Contrast Limited Adaptive Histogram Equalization (CLAHE) and Adaptive Histogram Equalization (AHE). The ROI were segmented using two techniques; semiautomatic segmentation and manual delineation. The intra and inter-observer variability was compared between ROI in mammogram images with and without contrast enhancement using semiautomatic segmentation and manual delineation respectively.

2. Part II: By using techniques for contrast enhancement and segmentation that result in lower intra and inter-observer variability, 378 mammographic image, with 147 image labeled as benign and another 231 labeled as malignant. Two techniques of optimization were adopted; TPOT and Grid Search Algorithm that include 3 types of classifier; Support Vector Machine (SVM), Naive Bayes (NB),and Multi-Layer Perceptron-Artificial Neural Network (MLP-ANN).

1.6 Thesis Outline

This thesis consists of five chapters which will cover from Chapter 1 to Chapter 5. Chapter 1 contains research background. This includes problem statement, significance of study and objectives of study.

Chapter 2 includes literature review which provides the background informationregarding breast cancer. It addresses the two main types of breast cancer, including benign and malignant. Throughout this chapter, the key techniques and algorithms that are used in this research to develop the computer-aided diagnostic system are highlighted and explained. This chapter also presents a survey of existing studies on computer-based diagnostic systems for breast cancer detection. These studies coveR all main components of such systems such as segmentation, feature

extraction, feature selection and classification.

Chapter 3 describes the design of breast cancer diagnosis using two optimization techniques, TPOT and Grid Search algorithm. First, the design of a proposed approach is introduced. The requirements of image acquisition are then explained. The requirements of image processing and image segmentation are also discussed, followed by the feature extraction and feature selection processes. The chapter further elaborates on the requirements for classification of breast cancer. Finally, the performance of measurements used to evaluate two optimization techniques, TPOT and GS algorithm.

Chapter 4 This chapter presents the discussion and the results of the experiments carried out. The chapter demonstrates how the results of the proposed approach resolve the problems mentioned in the problem statements.

Chapter 5 concludes and summarizes the research contributions made. The achievements and objectives of the research with respect to the experimental results obtained are highlighted along with the key findings and significance of the research.

REFERENCES

Acharya, U. R., Koh, J. E. W., Hagiwara, Y., Tan, J. H., Gertych, A., Vijayananthan, A., ... & Yeong, C. H. (2018). Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features. *Computers in biology and medicine*, *94*, 11-18.

Akila, K., Jayashree, L. S., & Vasuki, A. (2015). Mammographic image enhancement using indirect contrast enhancement techniques–a comparative study. *Procedia Computer Science*, *47*, 255-261.

Akilandeswari, U., Nithya, R., & Santhi, B. (2012). Review on feature extraction methods in pattern classification. *European Journal of Scientific Research*, *71*(2), 265-272.

Ala'a, R., Jalab, H. A., Shivakumara, P., Ibrahim, R. W., & Obaidellah, U. H. (2020). Kidney segmentation in MR images using active contour model driven by fractional-based energy minimization. *Signal, Image and Video Processing*, *14*(7),1361-1368.

Al-Najdawi, N., Biltawi, M., & Tedmori, S. (2015). Mammogram image visual enhancement, mass segmentation and classification. *Applied Soft Computing*, *35*, 175-185.

Al Rasyid, M. B., Arnia, F., & Munadi, K. (2018, February). Histogram statistics and GLCM features of breast thermograms for early cancer detection. In *2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON)* (pp. 120-124). IEEE.

Anitha, J., Peter, J. D., & Pandian, S. I. A. (2017). A dual stage adaptive thresholding (DuSAT) for automatic mass detection in mammograms. *Computer methods and programs in biomedicine*, *138*, 93-104.

Arslan, S., Ozyurek, E., & Gunduz-Demir, C. (2014). A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images. *Cytometry Part A*, *85*(6), 480-490.

Asmare, M. H., Asirvadam, V. S., & Hani, A. F. M. (2015). Image enhancement based on contourlet transform. *Signal, Image and Video Processing*, *9*(7), 1679-1690.

Assegie, T. A. (2021). An optimized K-Nearest Neighbor based breast cancer detection. *Journal of Robotics and Control (JRC)*, *2*(3), 115-118.

Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., ...

&

Zanetti, G. (2016). Making sense of big data in health research: towards an EUaction plan. *Genome medicine*, *8*(1), 1-13.

Azam, S., Eriksson, M., Sjölander, A., Gabrielson, M., Hellgren, R., Czene, K., & Hall, P. (2021). Mammographic microcalcifications and risk of breast cancer. *British Journal of Cancer*, 1-7.

Azizah, A. M., Hashimah, B., Nirmal, K., Siti Zubaidah, A. R., Puteri, N. A., Nabihah, A., ... & Azlina, A. A. (2019). Malaysia National cancer registry report (MNCR).

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Jama*, *319*(13), 1317-1318.

Berg, W. A., Gutierrez, L., NessAiver, M. S., Carter, W. B., Bhargavan, M., Lewis, R. S., & Ioffe, O. B. (2004). Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. Radiology, 233(3), 830-849.

Berment, H., Becette, V., Mohallem, M., Ferreira, F., & Chérel, P. (2014). Masses inmammography: What are the underlying anatomopathological lesions?. *Diagnostic and interventional imaging*, *95*(2), 124-133.

Birdwell, R. L. (2009). The preponderance of evidence supports computer-aided detection for screening mammography. *Radiology*, *253*(1), 9-16.

Birdwell, R. L., Ikeda, D. M., O'Shaughnessy, K. F., & Sickles, E. A. (2001). Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology*, *219*(1), 192-202.

Blanks, R. G., Wallis, M. G., & Moss, S. M. (1998). A comparison of cancer detection rates achieved by breast cancer screening programmes by number of readers, for one and two view mammography: results from the UK National Health Service breast screening programme. *Journal of Medical screening*, *5*(4), 195-201.

Bouaziz, S., Dhahri, H., Alimi, A. M., & Abraham, A. (2016). Evolving flexible beta basis function neural tree using extended genetic programming & hybrid artificial bee colony. *Applied Soft Computing*, *47*, 653-668.

Boyd, N. F., Guo, H., Martin, L. J., Sun, L., Stone, J., Fishell, E., ... & Yaffe, M. J. (2007). Mammographic density and the risk and detection of breast cancer. *New England journal of medicine*, *356*(3), 227-236.

Breast Cancer Treatment Options - Virginia Oncology. (2021). Retrieved 22 July 2021, from https://virginiacancer.com/breast-cancer/treatment-options/

Brynolfsson, P., Nilsson, D., Torheim, T., Asklund, T., Karlsson, C. T., Trygg, J., ... & Garpebring, A. (2017). Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters. *Scientific reports*, *7*(1), 1-11.

Burton, A., Byrnes, G., Stone, J., Tamimi, R. M., Heine, J., Vachon, C., ... & McCormack, V. A. (2016). Mammographic density assessed on paired raw and processed digital images and on paired screen-film and digital images across three mammography systems. *Breast cancer research*, *18*(1), 1-12.

Bustamam, A., Bachtiar, A., & Sarwinda, D. (2019). Selecting features subsets based on support vector machine-recursive features elimination and One Dimensional- Naïve Bayes classifier using support vector machines for classification of prostate and breast cancer. *Procedia Computer Science*, *157*, 450-458.

Chen, Q., Meng, Z., & Su, R. (2020). WERFE: A gene selection algorithm based on recursive feature elimination and ensemble strategy. *Frontiers in bioengineering and biotechnology*, *8*, 496

Chen, G., Xie, X., & Li, S. (2020). Research on Complex Classification Algorithm of Breast Cancer Chip Based on SVM-RFE Gene Feature Screening. *Complexity*, *2020*.

Chu, C., Hsu, A. L., Chou, K. H., Bandettini, P., Lin, C., & Alzheimer's Disease Neuroimaging Initiative. (2012). Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *Neuroimage*, *60*(1), 59-70.

Ciecholewski, M. (2017). Malignant and benign mass segmentation in mammograms using active contour methods. *symmetry*, *9*(11), 277.

Cooper, G. M., & Hausman, R. E. (2000). The development and causes of cancer. *Thecell: A molecular approach*, *2*.

Costa, E., Lorena, A., Carvalho, A. C. P. L. F., & Freitas, A. (2007, July). A review ofperformance evaluation measures for hierarchical classifiers. In *Evaluation methods for machine learning II: Papers from the AAAI-2007 workshop* (pp. 1-6).

Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology*, *3*(02),185-205.

Doi, K., MacMahon, H., Katsuragawa, S., Nishikawa, R. M., & Jiang, Y. (1999). Computer-aided diagnosis in radiology: potential and pitfalls. *European journal ofRadiology*, *31*(2), 97-109.

Doi, K. (2009). Computer-aided diagnosis in medical imaging: achievements and challenges. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany* (pp. 96-96). Springer, Berlin, Heidelberg.

Dorn, P. L., Al-Hallaq, H. A., Haq, F., Goldberg, M., Abe, H., Hasan, Y., & Chmura, S. J. (2013). A prospective study of the utility of magnetic resonance imaging in determining candidacy for partial breast irradiation. *International Journal of Radiation Oncology* Biology* Physics*, *85*(3), 615-622.

Eddy, D. M., Hasselblad, V., McGivney, W., & Hendee, W. (1988). The value of mammography screening in women under age 50 years. *Jama*, *259*(10), 1512-1519.

Ekpo, E. U., Alakhras, M., & Brennan, P. (2018). Errors in mammography cannot be solved through technology alone. *Asian Pacific journal of cancer prevention: APJCP*, *19*(2), 291.

Elshawi, R., Maher, M., & Sakr, S. (2019). Automated machine learning: State-of-the-art and open challenges. *arXiv preprint arXiv:1906.02287*.

Esposito, F., & Malerba, D. (2001). Machine learning in computer vision. *Applied Artificial Intelligence*, *15*(8), 693-705.

Fabijańska, A., & Sankowski, D. (2009). Computer vision system for high temperature measurements of surface properties. *Machine Vision and Applications*, *20*(6), 411-421.

Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, Ren, G. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes & diseases*, *5*(2), 77-106.

Flores, W. G., & de Albuquerque Pereira, W. C. (2017). A contrast enhancement method for improving the segmentation of breast lesions on ultrasonography. *Computers in biology and medicine*, *80*, 14-23.

Giger, M. L., Karssemeijer, N., & Armato, S. G. (2001). Computer-aided diagnosis in medical imaging.

Gordon, R., & Rangayyan, R. M. (1984). Feature enhancement of film mammograms using fixed and adaptive neighborhoods. *Applied optics*, *23*(4), 560-564.

Gowri, D. S., & Amudha, T. (2014, March). A review on mammogram image enhancement techniques for breast cancer detection. In *2014 International Conference on Intelligent Computing Applications* (pp. 47-51). IEEE.

Hamim, M., El Moudden, I., Moutachaouik, H., & Hain, M. (2020, June). Decision Tree Model Based Gene Selection and Classification for Breast Cancer Risk Prediction. In *International Conference on Smart Applications and Data Analysis* (pp. 165-177). Springer, Cham.

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621

Hemalatha, R. J., Thamizhvani, T. R., Dhivya, A. J. A., Joseph, J. E., Babu, B., & Chandrasekaran, R. (2018). Active contour based segmentation techniques for medical image analysis. *Medical and Biological Image Analysis*, *4*, 17.

Houssami, N., Turner, R., & Morrow, M. (2013). Preoperative magnetic resonance imaging in breast cancer: meta-analysis of surgical outcomes. *Annals of surgery*, *257*(2), 249-255.

Hiary, H., Zaghloul, R., Al-Adwan, A., & Moh'd B, A. Z. (2017). Image contrast enhancement using geometric mean filter. *Signal, Image and VideoProcessing*, *11*(5), 833-840.

Htay, T. T., & Maung, S. S. (2018, September). Early stage breast cancer detection system using glcm feature extraction and k-nearest neighbor (k-NN) on mammography image. In *2018 18th International Symposium on Communications and Information Technologies (ISCIT)* (pp. 171-175). IEEE.

Hubbard, R. A., Kerlikowske, K., Flowers, C. I., Yankaskas, B. C., Zhu, W., & Miglioretti, D. L. (2011). Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Annals of internal medicine*, *155*(8), 481-492.

Hwang, J., Kim, H. J., Lemaillet, P., Wabnitz, H., Grosenick, D., Yang, L., ... & Pogue, B. (2017, March). Polydimethylsiloxane tissue-mimicking phantoms for quantitative optical medical imaging standards. In *Design and Quality for Biomedical Technologies X* (Vol. 10056, p. 1005603). International Society for Optics and Photonics.

Jackins, V., Vimal, S., Kaliappan, M., & Lee, M. Y. (2021). AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *The Journal of Supercomputing*, *77*(5), 5198-5219.

Jadoon, M. M., Zhang, Q., Haq, I. U., Butt, S., & Jadoon, A. (2017). Three-class mammogram classification based on descriptive CNN features. *BioMed research international*, *2017*.

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(1), 4-37.

Jalalian, A., Mashohor, S., Mahmud, R., Karasfi, B., Saripan, M. I. B., & Ramli, A. R. B. (2017). Foundation and methodologies in computer-aided diagnosis systemsfor breast cancer detection. *EXCLI journal*, *16*, 113.

Jenifer, S., Parasuraman, S., & Kadirvelu, A. (2016). Contrast enhancement and brightness preserving of digital mammograms using fuzzy clipped contrast-limitedadaptive histogram equalization algorithm. *Applied Soft Computing*, *42*, 167-177.

Jeon, H., & Oh, S. (2020). Hybrid-Recursive Feature Elimination for Efficient FeatureSelection. *Applied Sciences*, *10*(9), 3211.

Jović, A., Brkić, K., & Bogunović, N. (2015, May). A review of feature selection methods with applications. In *2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO)* (pp. 1200-1205). Ieee.

Kim, S. E., Jeon, J. J., & Eom, I. K. (2016). Image contrast enhancement using entropy scaling in wavelet domain. *Signal Processing*, *127*, 1-11.

Kinnear, K. E., Langdon, W. B., Spector, L., Angeline, P. J., & O'Reilly, U. M.(Eds.). (1994). *Advances in genetic programming* (Vol. 3). MIT press.

Kumar, A. (2008). Computer-vision-based fabric defect detection: A survey. *IEEE transactions on industrial electronics*, *55*(1), 348-363.

Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B., ... & Gillies, R. J. (2012). Radiomics: the process and the challenges. *Magnetic resonance imaging*, *30*(9), 1234-1248.

Lai, D. T., Pakkanen, J., Begg, R., & Palaniswami, M. (2008). Computational Intelligence and Sensor Networks for Biomedical Systems. In *Encyclopedia of healthcare information systems* (pp. 261-273). IGI Global.

Laksmi, T. V., Madhu, T., Kavya, K., & Basha, S. E. (2016). Novel image enhancement technique using CLAHE and wavelet transforms. *International Journal of Scientific Engineering and Technology*, *5*(11), 507-511.

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., Van Stiphout, R. G., Granton, P., ... & Aerts, H. J. (2012). Radiomics: extracting more informationfrom medical images using advanced feature analysis. *European journal of cancer*, *48*(4), 441-446.

Larue, R. T., Defraene, G., De Ruysscher, D., Lambin, P., & Van Elmpt, W. (2017). Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *The British journal of radiology*, *90*(1070), 20160665.

Le, T. T., Fu, W., & Moore, J. H. (2020). Scaling tree-based automated machine learning to biomedical big data with a feature set selector. *Bioinformatics*, *36*(1), 250-256.

Lehman, C. D., Arao, R. F., Sprague, B. L., Lee, J. M., Buist, D. S., Kerlikowske, K., ... & Miglioretti, D. L. (2017). National performance benchmarks for modern screening digital mammography: update from the Breast Cancer Surveillance Consortium. *Radiology*, *283*(1), 49-58.

Lehman, C. D., Gatsonis, C., Kuhl, C. K., Hendrick, R. E., Pisano, E. D., Hanna, L. & Schnall, M. D. (2007). MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer. *New England Journal of Medicine*, *356*(13), 1295-1303.

Love, S. M., & Barsky, S. H. (2004). Anatomy of the nipple and breast ducts revisited. *Cancer*, *101*(9), 1947-1957.

Liao, M., Zhou, J., Sale, M., & Xiao, J. J. (2020). Application of machine learningand grid search approaches to minimize lucitanib pharmacokinetic variability following different dosing regimens.

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid search, random search, genetic

algorithm: a big comparison for NAS. *arXiv preprint arXiv:1912.06059*.

Mahesh, B. (2020). Machine Learning Algorithms-A Review. *International Journalof Science and Research (IJSR).[Internet]*, *9*, 381-386.

Mall, P. K., Singh, P. K., & Yadav, D. (2019, December). Glcm based feature extraction and medical x-ray image classification using machine learning techniques. In *2019 IEEE Conference on Information and Communication Technology* (pp. 1-6). IEEE.

Mattonen, S. A., Palma, D. A., Johnson, C., Louie, A. V., Landis, M., Rodrigues, G & Ward, A. D. (2016). Detection of local cancer recurrence after stereotactic ablative radiation therapy for lung cancer: physician performance versus radiomic assessment. *International Journal of Radiation Oncology* Biology* Physics*, *94*(5), 1121-1128.

Nabizadeh, N., & Kubat, M. (2015). Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. *Computers & Electrical Engineering*, *45*, 286-301.

Nayak, T., Bhat, N., Bhat, V., Shetty, S., Javed, M., & Nagabhushan, P. (2019). Automatic segmentation and breast density estimation for cancer detection usingan efficient watershed algorithm. In *Data Analytics and Learning* (pp. 347-358). Springer, Singapore.

Nelson, H. D., Pappas, M., Cantor, A., Griffin, J., Daeges, M., & Humphrey, L. (2016). Harms of breast cancer screening: systematic review to update the 2009US Preventive Services Task Force recommendation. *Annals of internal medicine*, *164*(4), 256-267.

Nikoo, H., Talebi, H., & Mirzaei, A. (2011, November). A supervised method for determining displacement of gray level co-occurrence matrix. In *2011 7th Iranian conference on machine vision and image processing* (pp. 1-5). IEEE.

O'Connor, J. P., Aboagye, E. O., Adams, J. E., Aerts, H. J., Barrington, S. F., Beer, A. J., ... & Waterton, J. C. (2017). Imaging biomarker roadmap for cancer studies. *Nature reviews Clinical oncology*, *14*(3), 169-186.

Oliver i Malagelada, A. (2007). *Automatic mass segmentation in mammographic images*. Universitat de Girona.

Olson, R. S., & Moore, J. H. (2016, December). TPOT: A tree-based pipeline optimization tool for automating machine learning. In *Workshop on automatic machine learning* (pp. 66-74). PMLR.

Orlenko, A., Kofink, D., Lyytikäinen, L. P., Nikus, K., Mishra, P., Kuukasjärvi, P., ... & Moore, J. H. (2020). Model selection for metabolomics: predicting diagnosis of coronary artery disease using automated machine learning. *Bioinformatics*, *36*(6),

1772-1778.

Osowski, S., Siroic, R., Markiewicz, T., & Siwek, K. (2008). Application of support vector machine and genetic algorithm for improved blood cell recognition. *IEEE Transactions on Instrumentation and Measurement*, *58*(7), 2159-2168.

Padhi, S., Rup, S., Saxena, S., & Mohanty, F. (2019, September). Mammogram Segmentation Methods: A Brief Review. In *2019 2nd International Conference onIntelligent Communication and Computational Techniques (ICCT)* (pp. 218-223). IEEE.

Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert review of precision medicine and drug development*, *1*(2), 207- 226.

Parmar, C., Rios Velazquez, E., Leijenaar, R., Jermoumi, M., Carvalho, S., Mak, R. H., ... & Aerts, H. J. (2014). Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PloS one*, *9*(7), e102107.

Pasha, S. S., Babu, P. S., & Vakil, Z. (2019, September). Enhancement of MRI Brain Images with Histogram Equalization Techniques. In *2019 International Conference on Emerging Trends in Science and Engineering (ICESE)* (Vol. 1, pp. 1-4). IEEE.

Patil, P. S., & Pawade, P. P. (2016). Biomedical image brightness preservation and segmentation technique using CLAHE and Wiener filtering. *Int. J. Adv. Res. Comput. Commun. Eng.*, *5*(4), 808-812.

Phi, X. A., Tagliafico, A., Houssami, N., Greuter, M. J., & de Bock, G. H. (2018). Digital breast tomosynthesis for breast cancer screening and diagnosis in women with dense breasts–a systematic review and meta-analysis. *BMC cancer*, *18*(1), 1- 9.

Pinsky, R. W., & Helvie, M. A. (2010). Mammographic breast density: effect on imaging and breast cancer risk. *Journal of the National Comprehensive Cancer Network*, *8*(10), 1157-1165.

ping Tian, D. (2013). A review on image feature extraction and representation techniques. *International Journal of Multimedia and Ubiquitous Engineering*, *8*(4), 385-396.

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ... & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, *39*(3), 355-368.

Punithavathy, K., Ramya, M. M., & Poobal, S. (2015, February). Analysis of statistical texture features for automatic lung cancer detection in PET/CT images. In *2015 International Conference on Robotics, Automation, Control and*

Embedded Systems (RACE) (pp. 1-5). IEEE.

Qayyum, A., & Basit, A. (2016, October). Automatic breast segmentation and cancer detection via SVM in mammograms. In *2016 International conference on emerging technologies (ICET)* (pp. 1-6). IEEE.

Qiu, Q., Duan, J., Gong, G., Lu, Y., Li, D., Lu, J., & Yin, Y. (2017). Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma. *Transl Cancer Res*, *6*(5), 940-948.

Raj, P., & David, P. E. (2020). *The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases*. Academic Press.

Raj, S. D., Shurafa, M., Shah, Z., Raj, K. M., Fishman, M. D., & Dialani, V. M. (2019). Primary and secondary breast lymphoma: clinical, pathologic, and multimodality imaging review. Radiographics, 39(3), 610-625.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347-1358.

Rangayyan, R. M., Shen, L., Shen, Y., Desautels, J. L., Bryant, H., Terry, T. J., ... & Rose, M. S. (1997). Improvement of sensitivity of breast cancer diagnosis

with adaptive neighborhood contrast enhancement of mammograms. *IEEE transactionson information technology in biomedicine*, *1*(3), 161-170.

Rangayyan, R. M., Ayres, F. J., & Desautels, J. L. (2007). A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, *344*(3-4), 312-348

Reddy, V. N., & Rao, P. S. (2018). Comparative analysis of breast cancer detection using K-means and FCM & EM segmentation techniques. *Ingénierie des Systèmesd'Information*, *23*(6).

Sakai, A., Onishi, Y., Matsui, M., Adachi, H., Teramoto, A., Saito, K., & Fujita, H. (2020). A method for the automated classification of benign and malignant masses on digital breast tomosynthesis images using machine learning and radiomic features. *Radiological physics and technology*, *13*(1), 27-36.

Saleem, A., Beghdadi, A., & Boashash, B. (2012). Image fusion-based contrast enhancement. *EURASIP Journal on Image and Video Processing*, *2012*(1), 1-17.

Santhi, K., & Banu, R. W. (2014). Contrast enhancement using brightness preserving histogram plateau limit technique.

Seal, A., Bhattacharjee, D., & Nasipuri, M. (2018). Predictive and probabilistic model for cancer detection using computer tomography images. *Multimedia Tools and Applications*, *77*(3), 3991-4010.

Shekar, B. H., & Dagnew, G. (2019, February). Grid search-based hyperparameter tuning and classification of microarray cancer data. In *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)* (pp. 1-8). IEEE.

Shi, H., & Liu, Y. (2011, September). Naïve Bayes vs. support vector machine: resilience to missing data. In *International Conference on Artificial Intelligence and Computational Intelligence* (pp. 680-687). Springer, Berlin, Heidelberg.

Sim, J. A., Kim, Y. A., Kim, J. H., Lee, J. M., Kim, M. S., Shim, Y. M., ... & Yun, Y.
H. (2020). The major effects of health-related quality of life on 5-year survival prediction among lung cancer survivors: applications of machine learning. *Scientific reports*, *10*(1), 1-12.

Simos, D., Catley, C., van Walraven, C., Arnaout, A., Booth, C. M., McInnes, M., ...& Clemons, M. (2015). Imaging for distant metastases in women with early-stage breast cancer: a population-based cohort study. *CMAJ*, *187*(12), E387-E397.

Singh, N. P., & Srivastava, R. (2016). Retinal blood vessels segmentation by using Gumbel probability distribution function based matched filter. *Computer methods and programs in biomedicine*, *129*, 40-50.

Skaane, P., Engedal, K., & Skjennald, A. (1997). Interobserver variation in the interpretation of breast imaging: comparison of mammography, ultrasonography, and both combined in the interpretation of palpable noncalcified breast masses. *Acta Radiologica*, *38*(4), 497-502.

Souza, J. C., Silva, T. F. B., Rocha, S. V., Paiva, A. C., Braz, G., Almeida, J. D. S., & Silva, A. C. (2017). Classification of Malignant and Benign tissues in mammography using dental shape descriptors and shape distribution.

Sovierzoski, M. A., Schwarz, L., & de Azevedo, F. M. (2009). Evaluation of benchmark indexes to determine the best performance of a binary neural classifier.In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany* (pp. 464-467). Springer, Berlin, Heidelberg.

Su, X., Chen, N., Sun, H., Liu, Y., Yang, X., Wang, W., ... & Yue, Q. (2020).
Automated machine learning based on radiomics features predicts H3 K27M mutation in midline gliomas of the brain. *Neuro-oncology*, *22*(3), 393-401.

Sundaram, M., Ramar, K., Arumugam, N., & Prabin, G. (2011). Histogram modified local contrast enhancement for mammogram images. *Applied soft computing*, *11*(8), 5809-5816.

Vachon, C. M., Van Gils, C. H., Sellers, T. A., Ghosh, K., Pruthi, S., Brandt, K. R., & Pankratz, V. S. (2007). Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Research*, *9*(6), 1-9.

Vicente, A. G., Castrejón, A. S., Amo-Salas, M., Fidalgo, J. L., Sanchez, M. M., Cabellos, R. A., ... & Madero, V. M. (2016). Glycolytic activity in breast cancer using 18F-FDG PET/CT as prognostic predictor: A molecular phenotype approach. *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)*, *35*(3), 152-158.

Vyborny, C. J., Giger, M. L., & Nishikawa, R. M. (2000). Computer-aided detection and diagnosis of breast cancer. *Radiologic Clinics of North America*, *38*(4), 725- 740.

Wang, Y., Aghaei, F., Zarafshani, A., Qiu, Y., Qian, W., & Zheng, B. (2017). Computer-aided classification of mammographic masses using visually sensitive image features. *Journal of X-ray Science and technology*, *25*(1), 171-186.

Watomakin, D. B., & Emanuel, A. W. R. (2019, October). Comparison of Performance Support Vector Machine Algorithm and Naive Bayes for Diabetes Diagnosis. In *2019 5th International Conference on Science in Information Technology (ICSITech)* (pp. 89-94). IEEE.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), 67-82.

Woods, B. J., Clymer, B. D., Kurc, T., Heverhagen, J. T., Stevens, R., Orsdemir,A., ... & Knopp, M. V. (2007). Malignant-lesion segmentation using 4D co-occurrence texture analysis applied to dynamic contrast-enhanced magnetic resonance breast image data. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *25*(3), 495-501.

Wottschel, V., Chard, D. T., Enzinger, C., Filippi, M., Frederiksen, J. L., Gasperini, C., ... & Ciccarelli, O. (2019). SVM recursive feature elimination analyses of structural brain MRI predicts near-term relapses in patients with clinically isolated syndromes suggestive of multiple sclerosis. *NeuroImage: Clinical*, *24*, 102011.

Zhang, H. (2015). Microwave imaging for ultra-wideband antenna based cancer detection.

Zhang, X. W., Song, J. Q., Lyu, M. R., & Cai, S. J. (2004). Extraction of karyocytes and their components from microscopic bone marrow images based on regional color features. *Pattern Recognition*, *37*(2), 351-361.

Zhang, Y., Deng, Q., Liang, W., & Zou, X. (2018). An efficient feature selection strategy based on multiple support vector machine technology with gene expression data. *BioMed research international*, *2018*.

Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. *Graphics gems*, 474-485.