Executive Summary of the thesis entitled

Identification, Analysis and Control of Breast Cancer using Mathematical Techniques

---

Submitted by

Dave Drashti Krutarth

(Registration No.: FOTE/1026)

Guided by

Dr. Trupti Pankaj Shah

Towards the Partial Fulfilment for the

Degree of

DOCTOR OF PHILOSOPHY

In

APPLIED MATHEMATICS

DEPARTMENT OF APPLIED MATHEMATICS

Faculty of Technology & Engineering

The Maharaja Sayajirao University of Baroda

Vadodara-390001, Gujarat, India

Table of Contents of the Thesis

Table of Contents of the Executive Summary

Introduction:

Cancer, a multifaceted and potentially fatal disease, involves the uncontrolled proliferation of cells that can metastasize, spreading to other body parts. With over 200 distinct types, the most prevalent globally include lung, breast, colorectal, prostate, and stomach cancers. In India, breast, oral, cervical, stomach, and lung cancers dominate, according to GLOBOCAN 2018. Statistics indicate approximately 2.25 million cases in India, with an alarming increase predicted.

Breast cancer, characterized by the growth of malignant cells in breast tissues, is the most common cancer in Indian women, representing 14% of all female cancers. It often manifests as lumps within the breast and can metastasize if not detected early. The disease's incidence rises from the early thirties and peaks between 50-64 years.

At the cellular level, mutations in genes controlling cell growth trigger cancer. Breast cancer typically develops in the lobules or ducts, potentially spreading to lymph nodes and beyond. It is categorized into invasive and non-invasive types, with invasive ductal carcinoma being the most common. The specific type influences treatment options and outcomes.

Metastatic breast cancer, or stage 4, has advanced beyond the breast to other body parts. Treatment plans, crafted by oncologists, aim to inhibit tumor growth and may include surgery, chemotherapy, radiation, targeted therapy, hormone therapy, and dietary modifications, like the ketogenic diet. However, treatments often bring side effects like hair loss, nausea, and fatigue due to their inability to distinguish between healthy and cancerous cells.

Recent studies have explored dietary strategies, such as the ketogenic diet, for cancer treatment. This diet, high in fats and low in carbohydrates, alters the body's energy metabolism, potentially impacting cancer cell growth.

In summary, cancer remains a formidable health challenge worldwide, with breast cancer being a significant concern in India. Advances in treatment and dietary approaches offer hope, but the disease's complexity necessitates continued research and personalized care strategies.

Importance:

Mathematical modeling in breast cancer research is essential, providing insights that revolutionize disease understanding and management. It dissects complex biological mechanisms, highlighting key aspects of tumor growth and metastasis. These models are fundamental in predicting disease progression, pivotal for early diagnosis and devising effective treatments. Particularly in treatment planning, they enable the simulation of various therapeutic scenarios, helping to tailor treatments to individual patients' needs. This approach enhances the efficacy of existing treatments and aids in developing new ones. By predicting responses to different therapies, mathematical models facilitate more personalized, targeted treatments, reducing side effects and improving patient outcomes. They also play a crucial role in optimizing treatment dosages and schedules, maximizing therapeutic efficiency while minimizing harm. In summary, mathematical modeling is invaluable in breast cancer research, significantly advancing treatment strategies and personalizing patient care.

The utilization of applied mathematics in medical science has witnessed significant growth, with a notable emphasis on the diagnosis of breast cancer using various soft computing methodologies. This thesis delves into the application of kernel-based approaches for the accurate identification of breast cancer, illustrating the expanding role of soft computing in medical research. The integration of Artificial Intelligence (AI), Machine Learning (ML), and Data Mining techniques is revolutionizing the way medical professionals approach cancer treatment. These technologies assist in navigating the plethora of treatment options, harnessing data from diverse databases to pinpoint optimal drug choices tailored to individual patients.

AI's contribution extends to enhancing decision-making in drug selection, broadening treatment applications, and streamlining clinical trials by efficiently identifying suitable patients from extensive data pools. The term 'AI in healthcare' encompasses the application of ML algorithms and software tools, simulating human cognitive abilities in the analysis, presentation, and interpretation of complex medical data. A key function of AI in this realm is classification, a data mining technique critical for early disease detection and analysis, potentially lifesaving in medical treatment and bioinformatics.

Further, AI is increasingly instrumental in various aspects of patient care and the development of intelligent health systems. Research indicates that AI can match or surpass human performance in crucial healthcare tasks, such as disease diagnosis. The scope of AI techniques in healthcare is vast, ranging from machine learning to deep learning, and plays a pivotal role in areas like disease diagnosis, drug discovery, and patient risk assessment. For accurate disease diagnosis using AI, an array of medical data sources is essential, including ultrasound, magnetic resonance imaging, mammography, genomics, and computed tomography scans, among others. This comprehensive approach exemplifies the transformative impact of AI and related technologies in the medical field, particularly in breast cancer diagnosis and treatment.

Literature survey:

In the extensive field of breast cancer research and computational methods, numerous studies and resources have significantly advanced our understanding and treatment approaches. Healthline's comprehensive guide provided a foundational understanding of the disease, covering symptoms, stages, and types[1]. Emphasizing the critical importance of awareness and early detection, Firstpost's 2019 article marked a significant contribution to the public discourse on cancer[2]. Similarly, the National Health Portal of India's initiative for Breast Cancer Awareness Month in 2019 played a pivotal role in educating and raising awareness among the public regarding breast cancer[3].

The University of California, Irvine's machine-learning database, particularly the breast-cancer-wisconsin dataset, emerged as a crucial resource for computational studies, offering extensive data for researchers[4]. Additionally, the cBioPortal for Cancer Genomics provided an extensive database for genomic studies, aiding significantly in advanced cancer research[5]. In the theoretical domain, Eduardo Sontag's "Mathematical Control Theory" offered essential frameworks for understanding cancer growth dynamics and treatment responses[6].

Several groundbreaking studies have also contributed to the field. In 2014, Allen et al. explored ketogenic diets as a potential adjuvant cancer therapy, examining its history and mechanisms[7]. The 2002 study by Pinho, Freedman, and Nani developed a chemotherapy model for cancer with metastasis, making a significant impact on mathematical modeling in oncology[8]. D'Onofrio et al.'s 2009 research focused on the optimal delivery of combination therapy for tumors, adding valuable insights to treatment strategies[9].

Further, the 2005 study by De Pillis, Radunskaya, and Wiseman validated a mathematical model of cell-mediated immune response to tumor growth, offering a new perspective on cancer treatment[10]. Mufudza et al.'s 2012 research assessed the effects of estrogen on breast cancer dynamics, combining computational methods with medical insights[7]. Additionally, Abernathy et al.'s 2017 study on the global dynamics of a breast cancer competition model added a new dimension to our understanding of cancer behavior[11].

The integration of machine learning and artificial intelligence in breast cancer research has been notably showcased in studies like Isaac Oke, Matadi, and Xulu's 2018 optimal control analysis of a breast cancer model[12], and Maglogiannis et al.'s 2009 development of an SVM-based automated diagnosis system[13]. The advancements in AI and ML are further evident in Chen et al.'s 2011 work on SVM classifiers for breast cancer diagnosis[14], Ahmad et al.'s 2013 study on using machine learning techniques for predicting breast cancer recurrence[15], and Cruz-Roa et al.'s 2017 research employing deep learning for invasive breast cancer detection in whole-slide images[16].

The impact of big data in breast cancer research was discussed by Ibnouhsein et al. in 2018, highlighting the transformative role of data in this domain[17]. Foundational knowledge in neural networks and deep learning, essential for medical image analysis and diagnostics, has been provided by resources like the University of Wisconsin, Madison and various deep learning tutorials and surveys [19]. This collective body of work reflects the dynamic and interdisciplinary nature of breast cancer research, driven by rapid advancements in computational and mathematical methodologies.

Brief about our work:

In this research, we present a multifaceted approach to understanding and managing breast cancer, leveraging both mathematical modelling and advanced artificial intelligence techniques, with a specific focus on the BreakHis histopathological image dataset.

Our study begins with a mathematical model that encapsulates the dynamics of breast cancer, particularly focusing on the effectiveness of combined chemotherapy and monoclonal antibody treatment. Through the Z-control technique and extensive stability analysis, we demonstrate a significant reduction in cancer cell growth, confirmed by numerical computations and graphical representations.

We then compare two control strategies—Z-control and optimal control—applied to the same system, providing a thorough comparative analysis. This is complemented by numerical computations, enhancing our understanding of breast cancer dynamics control.

A major part of our research involves the application of artificial neural networks (ANNs) for the early detection and diagnosis of breast cancer, utilizing the BreakHis dataset. We introduce a novel weight updating algorithm for training these networks, showing superior performance over traditional methods like Gradient Descent and RMSprop in terms of accuracy and other metrics.

Additionally, we propose the Modified Firefly Algorithm to optimize neural network training for breast cancer detection, benchmarked against other nature-inspired algorithms. We employ advanced feature extraction techniques like Gabor filters and Local Binary Patterns, significantly enhancing our model's efficiency.

Finally, our research includes the use of a Convolutional Neural Network (CNN) for feature extraction from the BreakHis dataset, paired with our novel weight updating algorithm. This combination yields an unprecedented classification accuracy of 100%, showcasing its potential in early and accurate detection of malignant breast tissue.

In summary, our work provides a comprehensive and detailed exploration of breast cancer dynamics, utilizing mathematical models and cutting-edge AI techniques, all centered around the extensive and crucial BreakHis histopathological image dataset.

Organization of the thesis:

The layout of the thesis along with the proposed methodologies used in constructing the predictive models for classification of breast cancer (described in chapters 3 to 6) is as follows.

Chapter 1: Introduction

This chapter mainly deals with the motivation as well as literature survey of the breast cancer.

Chapter 2: Mathematical preliminaries

Chapter 3: Exploring the Dynamics of Breast Cancer Treatment: Comparative Analysis of Z-Control and Optimal Control Mathematical Models

This chapter presents a meticulous comparative analysis of two sophisticated mathematical models, each unravelling the intricate dynamics of breast cancer and the interplay among cancer cells, normal cells, and diverse treatment modalities. These modalities encompass chemotherapy, monoclonal antibody drugs, and the ketogenic diet, offering a comprehensive view of current therapeutic approaches.

The first part of the analysis delves into a model employing the Z-control technique, providing a profound stability analysis and illustrating its potential in significantly mitigating cancer cell proliferation. This model sets the stage for a nuanced understanding of cancer cell dynamics and the impact of various treatment strategies.

Building on this foundation, the second part of the chapter introduces a model that incorporates optimal control theory. This advanced model not only continues the narrative of cancer cell reduction but also accentuates the heightened effectiveness of combining multiple treatment modalities. The model serves as a testament to the potential synergies that can be achieved through an integrated treatment approach.

To bridge the gap between theoretical constructs and practical applicability, the chapter presents numerical simulations that validate the models' predictions. These simulations reinforce the robustness of the models and their relevance in real-world scenarios, paving the way for their potential use in precision medicine.

Collectively, the models explored in this chapter pave the way for a promising future in cancer therapy. They underscore the significance of adopting a multi-faceted and precision-driven approach to treatment, which holds the promise of improving patient outcomes and advancing the field of oncology.

Chapter 4: Enhancing Diagnostic Precision in Breast Cancer Histopathology: A Novel Weight Optimization Approach for ANN

This chapter critically addresses the grave public health concern posed by breast cancer, which predominantly impacts the global female demographic. It acknowledges the fundamental significance of histopathological examination in the early and accurate detection of this condition. Recent advancements in artificial neural networks (ANN) have shown promising potential in autonomously classifying histopathological images related to breast cancer. However, it is vital to recognize that the performance of such neural networks is significantly influenced by the optimization algorithm employed during their training phase.

Within the scope of this research, we introduce a ground breaking weight optimization algorithm. The study is meticulously designed around the development and training of an ANN model, leveraging this innovative algorithm specifically for breast cancer diagnosis. The efficiency of the newly developed ANN is critically assessed against models trained with established optimization techniques, including Gradient Descent (GD), Root Mean Square

Propagation (Rmsprop), and Gradient Descent with Momentum (GDM).

The research employs the Breakhis dataset, a pivotal resource in breast cancer histopathology. The initial stage involves the extraction of critical features from the dataset using techniques such as Gabor filters, Local Binary Patterns (LBP), and the Co-occurrence matrix. These features are then processed through the intricate layers of neural network models.

The empirical findings of this study are remarkable. The novel weight optimization algorithm not only achieves an unparalleled classification accuracy of 100% but also excels across various other key performance metrics such as the F1-score, precision, sensitivity, and specificity. These results decisively highlight the algorithm's profound efficacy in accurately distinguishing malignant from non-malignant breast tissue, marking a significant advancement in the realm of medical diagnostics.

Chapter 5: Advancements in Breast Cancer Diagnosis: Efficacy of the Modified Firefly Algorithm in Optimizing ANN-Based Histopathological Analysis

This chapter delves into the pressing issue of breast cancer, emphasizing the critical importance of early detection for successful treatment outcomes. It presents an innovative approach to refining the diagnostic accuracy of breast cancer through microscopic tissue analysis. Central to this study is the introduction of a Modified Firefly Algorithm, specifically designed to enhance the training efficacy of Artificial Neural Networks (ANNs). This novel approach is meticulously evaluated against conventional nature-inspired algorithms within the context of the BreakHis dataset, known for its comprehensive compilation of microscopic breast tissue images.

The methodological framework of the study involves a systematic pre-processing of the BreakHis dataset images, followed by a robust feature extraction process employing Gabor filters, Local Binary Patterns, and Co-occurrence matrices. These extracted features serve as inputs for the ANNs. The focal point of the research is the optimization of these networks through the implementation of the Modified Firefly Algorithm, with its performance critically compared against established optimization techniques including Firefly Optimization, Bat Optimization, Grey Wolf Optimization, and Particle Swarm Optimization.

The findings of the study are significant, with the Modified Firefly Algorithm displaying a marked superiority in optimizing ANNs for the nuanced task of breast cancer detection. When juxtaposed with traditional algorithms, it not only exhibited heightened accuracy in classifying breast tissue images as malignant or benign but also demonstrated exceptional proficiency across essential metrics such as the F1-score, precision, sensitivity, and specificity. This underlines its capability in the precise identification of detrimental breast tissue.

Conclusively, the Modified Firefly Algorithm emerges as a pivotal advancement in the training of ANNs, significantly bolstering the process of breast cancer detection through microscopic image analysis. Its exemplary performance across multiple evaluative metrics positions it as a formidable technique in the field of medical imaging. Prospective research avenues might involve exploring the applicability of this algorithm in diverse medical imaging scenarios, potentially amplifying its transformative impact on healthcare diagnostics.

Chapter 6: Optimizing Breast Cancer Detection: Integrating Convolutional Neural Networks and a Novel Weight Updating Algorithm in Histopathological Analysis

This chapter delves into the critical public health issue of breast carcinoma, emphasizing its global impact on women's health. It acknowledges the indispensable role of microscopic breast tissue analysis in the early stages of cancer detection and diagnosis. The study leverages the prowess of convolutional neural networks (CNN) for the advanced feature extraction from the BreakHis histopathology breast cancer dataset, thereby facilitating a comprehensive analysis of each image for the identification of crucial cancer markers.

Further, the chapter meticulously compares the performance of an Artificial Neural Network (ANN) model, enhanced with a novel weight updating algorithm, against models employing traditional optimization techniques such as Gradient Descent (GD), Root Mean Square Propagation Method (Rmsprop), and Gradient Descent with Momentum (GDM). The results obtained from this comparison are profound, with the ANN model showcasing an unparalleled classification accuracy of 100%. Moreover, the model exhibits superior performance in critical metrics, including F1-score, precision, sensitivity, and specificity, thereby demonstrating its effectiveness in the precise identification of malignant breast tissue.

The findings presented in this chapter underscore the efficiency and transformative potential of integrating CNN with an innovative weight updating algorithm, marking a significant advancement in the field of medical diagnostics and the early detection of breast cancer.

Key Findings

This research presents a multidisciplinary approach to breast cancer detection and management, integrating mathematical modeling, control strategies, and artificial intelligence techniques. The key findings of this study are as follows:

1. **Mathematical Modelling and Control Strategies:**
    - A mathematical model was developed to analyse breast cancer dynamics, incorporating chemotherapy and monoclonal antibody treatments.
    - Stability analysis, performed using the Lyapunov function, demonstrated the

impact of treatment interventions on tumor suppression.

- The effectiveness of control strategies was evaluated, revealing that the optimal control approach significantly outperforms the Z-control method in minimizing cancer cell proliferation.

2. **Artificial Intelligence-Based Classification:**
    - Convolutional Neural Networks (CNNs) were utilized for feature extraction, while Artificial Neural Networks (ANNs) were employed for the automated classification of histopathological breast cancer images from the BreakHis dataset.
    - Advanced feature extraction techniques, including Gabor filters, Local Binary Patterns (LBP), and the Co-occurrence matrix, enhanced classification performance.
    - A novel weight-updating optimization algorithm was introduced for ANN training, achieving superior classification accuracy compared to conventional optimization techniques such as Gradient Descent (GD), Root Mean Square Propagation (Rmsprop), and Gradient Descent with Momentum (GDM).
    - The proposed approach achieved 100% classification accuracy, excelling across key evaluation metrics, including F1-score, precision, sensitivity, and specificity.
3. **Optimization Strategies for Enhanced Performance:**
    - A Modified Firefly Algorithm was developed and compared with established nature-inspired optimization techniques, including Firefly, Bat, Grey Wolf, and Particle Swarm Optimization.
    - The Modified Firefly Algorithm exhibited superior proficiency in training neural networks, demonstrating high accuracy and robustness in classifying malignant breast tissue.

Conclusion:

The culmination of the research presented in these chapters offers a groundbreaking perspective on the management and diagnosis of breast cancer, a significant public health concern worldwide. This synopsis encapsulates the essence of multiple studies, each contributing uniquely to the field of oncology and precision medicine.

Chapter third laid the foundation by exploring the complex dynamics of breast cancer, utilizing mathematical models to examine the interplay between cancerous and normal cells under various treatment modalities. The introduction of the Z-control technique and subsequent stability analysis in this chapter provided a deep insight into the potential of mathematical modelling in reducing cancer cell growth. The progression to optimal control theory in the subsequent part of the chapter further enhanced our understanding, demonstrating the effectiveness of integrated treatment strategies. The numerical simulations presented validated these theoretical models, bridging the gap between abstract mathematical constructs and their practical medical implications.

In the next chapter, the focus shifted to the technological advancements in medical diagnostics, particularly the use of artificial neural networks (ANN) for breast cancer diagnosis. The development of a novel weight optimization algorithm and its application to ANN models represented a significant stride in this field. The rigorous comparative analysis against established optimization techniques, using the BreakHis dataset, highlighted the algorithm's remarkable ability to achieve 100% classification accuracy and excel in key performance metrics. This chapter underscored the profound impact of integrating cutting-edge computational techniques in enhancing the accuracy and efficiency of cancer diagnostics.

The fifth chapter presented an innovative approach to optimizing ANNs for breast cancer detection through the Modified Firefly Algorithm. The comprehensive analysis of this algorithm against traditional nature-inspired algorithms demonstrated its superiority in classifying breast tissue images. This chapter not only emphasized the algorithm's precision in diagnosis but also its potential applicability in various medical imaging domains.

Finally, the last chapter focused on the integration of convolutional neural networks (CNN) with the novel weight updating algorithm, further advancing the capabilities in breast cancer detection. The remarkable classification accuracy and performance in metrics like F1-score, precision, sensitivity, and specificity, highlighted in this chapter, affirmed the potential of such integrative approaches in medical diagnostics.

In conclusion, these chapters collectively provide a compelling narrative on the advancement of breast cancer diagnosis and treatment. From mathematical modeling to the application of sophisticated algorithms in ANN and CNN models, the research presented here opens up new avenues in precision medicine and cancer therapy. The consistent thread across these chapters is the pursuit of accuracy, efficiency, and effectiveness in breast cancer management, underscoring the importance of innovative approaches in combating this global health challenge. The insights gained from these studies pave the way for future research and development, potentially transforming the landscape of healthcare diagnostics and treatment strategies.

Future scope:

The future scope of our research, as delineated from the findings and methodologies of these studies, is expansive and multifaceted. There is significant potential for further development and refinement of the mathematical models and control strategies, particularly in the realm of personalized medicine, where individual patient data could be used to tailor treatments more effectively. The novel weight updating algorithm and the Modified Firefly Algorithm, demonstrated to be highly effective in classifying breast cancer histopathology images, offer promising avenues for application in other types of cancer detection and a broader range of medical imaging tasks. Additionally, the integration of these advanced algorithms with emerging technologies in AI and machine learning, such as deep learning and neural network fine-tuning, presents an exciting opportunity to enhance diagnostic accuracy and treatment efficacy. Furthermore, expanding the dataset to include a wider variety of cases and conditions could significantly improve the robustness and generalizability of these models, paving the way for their application in diverse clinical settings and populations.

Bibliography:

[1] https://www.healthline.com/health/breast-cancer

[2] https://www.firstpost.com/tech/science/world-cancer-day-2019

[3] https://www.nhp.gov.in/breast-cancer-awareness-month2019_pg

[4] www.archieve.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/.

[5] https://www.cbioportal.org

[6] Allen, B.G.; Bhatia, S.K.; Anderson, C.M.; Eichenberger-Gilmore, J.M.; Sibenaller, Z.A.; Mapuskar, K.A.; Schoenfeld, J.D.; Buatti, J.M.; Spitz, D.R.; Fath, M.A. , “Ketogenic diets as an adjuvant cancer therapy: History and potential mechanism”, Redox Biol. 2014, 2, 963–970.

[7] Mufudza, C.; Walter, S.; Chiyaka, E.T.,“Assessing the effects of estrogen on the dynamics

of breast cancer", Comput. Math. Methods Med. 2012,

[8] Pinho, S.T.R.D.; Freedman, H.I.; Nani, F.," A chemotherapy model for the treatment of cancer with metastasis", Math. Comput. Model. 2002, 36, 773–803.

[9] D'Onofrio, A.; Ledzewicz, U.; Maurer, H.; Schättler, H. , "On optimal delivery of combination therapy for tumors", Math. Biosci. 2009, 222, 13–26.

[10] De Pillis, L.G.; Radunskaya, A.E.; Wiseman, C.L. , "A validated mathematical model of cell-mediated immune response to tumor growth" , Cancer Res. 2005, 65, 7950–7958.

[11] Abernathy,K.;Abernathy, Z.;Baxter, A.;Stevens,M., "GlobalDynamicsofaBreastCancerCompetitionModel. Differ. Equ. Dyn. Syst. 2017, 3, 1–15.

[12] Segun Isaac Oke * ID , Maba Boniface Matadi and SibusisoSouthwellXulu , "Optimal Control Analysis of a Mathematical Model for Breast Cancer", Mathematical and Computational Applications, 2018, 23, 21; doi:10.3390/mca23020021

[13] IliasMaglogiannis, Elias Zafiropoulos, IoannisAnagnostopoulos, "An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers", Applied Intelligence, vol. 30, pg. 24-36 (2009)

[14] Hui-Ling Chen, Bo Yang, Jie Liu, Da-You Liu , "A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis" Expert systems with Applications, Vol. 38, pg. 9014-9022 (2011)

[15] Ahamad LG, Eshlaghy AT, Poorebrahimi A, Ebrahimi M , Razavi AR(2013), "Using Three Machine learning techniques for predicting Breast Cancer recurrence", Journal of Health and Medical Informatics , 4 :124. doi:10.4172/2157-7420.1000124

[16] Cruz-Roa, A. et al.," Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent" Sci. Rep. 7, 46450; doi: 10.1038/ srep46450 (2017).

[17] IssamIbnouhsein, Stéphane Jankowski, Karl Neuberger, and Carole Mathelin, "The Big

Data Revolution for Breast Cancer Patients" , Eur J Breast Health. 2018 Apr; 14(2): 61–62 (2018) doi: 10.5152/ejbh.2018.0101

[18] Benjamin C. Kuo, "Automatic Control Systems" Prentice Hall of India Private ltd.

[19] A Basic Introduction to Neural Network', University of Wisconsin, Madison. Available at: http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html