Chapter 1

Introduction

1.1 Introduction

Cancer is a multifaceted and potentially life-threatening disease characterized by the uncontrolled growth of cells, which may metastasize to different parts of the body. 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.

1.2 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 adoption of Artificial Intelligence (AI), Machine Learning (ML), and Data Mining methodologies is transforming the approach of medical professionals toward 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 concept of 'AI in healthcare' refers to the use of machine learning algorithms and software tools designed to emulate human cognitive functions for analyzing, presenting, and interpreting 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.

1.3 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 Sontags "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]. DOnofrio 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 [11]. 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 [12].

The application of machine learning and artificial intelligence in breast cancer research is exemplified by studies such as Isaac Oke, Matadi, and Xulus 2018 analysis, which focused on optimal control strategies for a breast cancer model [13], and Maglogiannis et al.'s 2009 development of an SVM-based automated diagnosis system [14]. The advancements in AI and ML are further evident in Chen et al.'s 2011 work on SVM classifiers for breast cancer diagnosis [15], Ahmad et al.s 2013 study on using machine learning techniques for predicting breast cancer recurrence [16], and The 2017 study by Cruz-Roa et al. utilized deep learning techniques for detecting invasive breast cancer in whole-slide imaging[17].

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 [18]. 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 [4]. This collective body of work reflects the dynamic and interdisciplinary nature of breast cancer research, driven by rapid advancements in computational and mathematical methodologies.

1.4 Brief About the 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 strategiesZ-control and optimal controlapplied to the same system, providing a thorough comparative analysis. This is complemented by numerical computations, enhancing our understanding of breast cancer dynamics control.

A significant focus of our research is the use of artificial neural networks (ANNs) for the early detection and diagnosis of breast cancer, leveraging 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.

Our research also incorporates the utilization of a Convolutional Neural Network (CNN) for feature extraction from the BreakHis dataset, combined with the implementation of a 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.

1.5 Organisation of Thesis

The structure of this thesis, along with the proposed methodologies employed in developing predictive models for breast cancer classification (outlined in Chapters 3 to 6), is presented as follows.

1.5.1 Chapter 1: Introduction

This chapter primarily focuses on the motivation behind the study and a comprehensive literature review of breast cancer.

1.5.2 Chapter 2: Mathematical Preliminaries

This chapter focuses on stability analysis in dynamical systems using Jacobian eigenvalues and explores optimal control theory through Z-control for error minimization. Key concepts like the Hamiltonian, Pontryagins Principle, and Python libraries for algorithm implementation are discussed. Machine learning evaluation metrics using confusion matrices are also addressed.

1.5.3 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 connect theoretical concepts with practical application, this chapter includes numerical simulations that confirm the accuracy of 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.

1.5.4 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 groundbreaking 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.

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

This chapter explores the critical issue of breast cancer, highlighting the paramount importance of early detection in achieving effective 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.

1.5.6 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 effec