Chapter 1
Analysis of Machine Learning Algorithms for Breast Cancer Detection

Aswathy M. A.
VIT University, India

Jagannath Mohan
https://orcid.org/0000-0001-8953-118X
VIT Chennai, India

ABSTRACT

As per the latest health ministry registries of 2017-2018, breast cancer among women has ranked number one in India and number two in United States. Despite the fact that breast cancer affects men also, pervasiveness is lower in men than women. This is the reason breast cancer is such a vital concern among ladies. Roughly 80% of cancer malignancies emerge from epithelial cells inside breast tissues. In breast cancer spectrum, ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC) are considered malignant cancers that need treatment and care. This chapter mainly deals with breast cancer and machine learning (ML) applications. All through this chapter, different issues related to breast cancer prognosis and early detection and diagnostic techniques using various ML algorithms are addressed.

INTRODUCTION

During the period of 2017-2018, health ministry is predicted that the breast cancer may reach 1797900 affected patients by 2020. An article in the journal consultant 360 says that every year approximately 200,000 or above new cases of breast cancer are reporting and 40,000 or above patients are dying with breast cancer in the United States (Estape, 2018). The incidence rate of breast cancer increases with the age of patients. Mostly 50% of new breast cancer diagnoses happening at an age of 65 years and older and the incidence of breast cancer rate increases till the 80s (Sharma, 2001). Usually pathologists detect breast cancer by manually adjusting region of interest and segmenting lesions from that selected area.

DOI: 10.4018/978-1-5225-9902-9.ch001

But chances of intra and inter observability variations will be there. In that context, it is very challenging to find a method that associates automatically selecting region of interests (ROIs) and differentiating ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC) from other normal cancers.

Traditionally cancer detection and the treatment were identified purely on pathologist's experience. These specialists were having above 15 years of practice in the medical field and have seen many patients in similar conditions. Still the accuracy was not 100%. With the evolution of machine learning and artificial intelligence, computer aided diagnosis of different types of cancer became easy. Machine intelligence (MI) or Artificial intelligence (AI) is defined as intelligence displayed by machines and adapt to the surroundings to perform actively to achieve its goals. Here the device imitates the "cognitive" functions correlated to human minds such as learning and perception. AI has exceptional impact in the field of image processing. Machine learning is a sub discipline of artificial intelligence and both are coming under computer science branch. The subject 'Machine learning' gained immense priority among pathologists and radiologists for its benefits in global health care industry and the curiosity boosted towards how the machine learning will strengthen medical specialists in their area of work. The study in machine learning makes healthcare engineering applications straightforward in acquiring and storing patient details in a database and possible to access these databases from anywhere in the world. Smart intelligent systems are there to help doctors and pathologists to interrogate and analyses these complex datasets. An article in Technology trends point out the world market of artificial intelligence including machine learning in various fields like medical imaging, health care, recognition and identification is estimated to reach 2 billion dollars in 2023 (Massat, 2018). Thus, the branch of artificial intelligence and machine learning became a valuable sphere in the healthcare industry.

Image processing became one of the fundamental elements in biomedical and medicinal research, laboratory areas etc. Image processing succeeds in processing a three-dimensional image and converts it into a two dimensional one using numerical data analysis. According to the last 7 years' market hype, artificial intelligence techniques like machine learning and other computer vision methods will change health care medical imaging industry in the stipulations like huge productivity, improved accuracy in diagnosis and good clinical outcomes (Massat, 2018). Hence machine learning would act as a mediator in between ever-growing figure of diagnostic image screening programs despite of harsh scarcity of radiologists and pathologists in many countries. Figure 1 shows the trend of using machine learning algorithms in the world market for various applications like computer vision and deep learning.

Among the vast variety of machine learning applications, cancer detection and prognosis diagnosis at its early stage is gaining high prominence in health care industry. Many researchers and specialists like IBM Watson had invested money and time in this field to make headway but lasted with little success in the history. Now Google came up with a machine learning system that incorporates microscope which helps doctors and pathologists in cancer detection. Similarly, there are many more algorithms that help in detecting any type of cancer with the help of machine learning and deep learning. Different AI techniques like fuzzy logic (FL), genetic algorithm, and neural network can be employed to solve many problems in the biomedical field. However, each method has its own constraints and can be used in certain circumstances only. In such situations, combination of these techniques might help in solving these problems. For example, neural network is just like a "black box" to the users. Users cannot access neural network interiors for understanding parameter revision. At the same time, fuzzy logic has issues in deciding membership functions (MFs). Artificial neural networks cannot parallelize and architecture selection is difficult. Genetic algorithm has an advantage of dealing complicated parameter optimization problems. Sometimes the combination of these techniques (hybrid systems) like neurofuzzy, neural