iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VI **Month of publication:** June 2023

DOI: https://doi.org/10.22214/ijraset.2023.53112

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VI Jun 2023- Available at www.ijraset.com

Review on Multiple Cancer Disease Prediction And Identification using Machine Learning Techniques

Bhushan S. Jakkanwar[1], Prof. Jayant Rohankar[2], Prof. Snehal Bagde[3]
[1]PG Scholar, [2]Project Guide, [3]Project Co-Guide, Artificail Intelliegence & Machine Learning, Tulsiramji Gaikwad Patil Collage of Engineering & Technology , Nagpur , India

Abstract: ***Cancer has been described as a diverse illness with a wide range of subgroups. Early cancer diagnosis and prognosis are essential for clinical patient treatment, which has become a requirement in cancer research. Numerous research teams from the biomedical and bioinformatics fields have studied the use of machine learning (ML) techniques due to the significance of categorising cancer patients into high or low risk groups. These methods have been applied in an effort to simulate the development and management of malignant diseases. Additionally, the importance of ML tools is demonstrated by their capacity to recognise important elements in complex datasets. Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs), and Decision Trees (DTs) are a few of the methods that have been widely used in cancer research to construct prediction models that enable precise and effective decision-making. Although it is clear that applying ML techniques can enhance our comprehension of how cancer progresses, more validation is required before these techniques can be used in routine clinical practise. Advance ML techniques are used in this work to mimic the detection of cancer in numerous human bodily organs, including the brain, breast, and lung.***

Keywords: Machine learning , Cancer susceptibility , Predictive models , Cancer recurrence , Cancer survival etc.

I. INTRODUCTION

Cancer research has seen a steady change throughout the past few decades. Researchers have used a variety of techniques, such as early-stage screening, to identify cancer types before symptoms appear. Additionally, they have created fresh methods for anticipating the results of cancer treatments early on. The medical research community has access to a lot of cancer data thanks to the development of new technologies in the field of medicine [1]. However, one of the most fascinating and difficult problems for doctors is making an accurate prediction of a disease's course. Consequently, ML techniques have gained popularity as a tool for medical researchers. These methods are capable of extracting patterns and correlations from large datasets, as well as efficiently forecasting the course of a particular cancer type [1][2].

We give a review of research that employ these techniques in relation to the prediction and prognosis of cancer in light of the significance of personalised medicine and the emerging trend on the application of ML techniques. In these research, prognostic and predictive factors that may be independent of a particular treatment or that are combined to suggest therapy for cancer patients, respectively, are taken into consideration. Additionally, we talk about the different ML techniques being utilised, the different kinds of data they incorporate, and the effectiveness of each proposed method, as well as their advantages and disadvantages [3].

The incorporation of mixed data, including clinical and genetic data, is a clear trend in the proposed works. Yet a recurring issue that we identified in numerous works is the absence of external validation or testing regarding the accuracy of their models. It is obvious that the use of ML techniques could increase the precision of predictions for cancer susceptibility, recurrence, and survival. Utilising ML approaches, the accuracy of cancer outcome prediction has increased by 15%–20% over the past few years [4].

Every person places "being healthy" at the top of their priority list. Everyone deserves to live a healthy life. And a lot of the time, our attitude is influenced by our feelings. everything, whether it's good or bad! The important thing to keep in mind in these circumstances is to maintain a "+ve Approach" to health [5].

Cancer is one such widespread illness in our environment. It is a deadly condition frequently brought on by the accumulation of hereditary disorders and many pathological alterations. Cancerous cells are aberrant, life-threatening growths that can appear anywhere on the human body. Therefore, the main goals of our project are to develop a platform for cancer diagnosis and improve the methods now in use. We are working to develop a web application that will benefit every cancer survivor. With the help of this programme, users will learn about common cancer types, their symptoms, dos and don'ts for treating each type, fund-related challenges, and other real-world problems. Motivational forum focusing on the latest ML techniques and DIP for the identification of breast cancer and brain tumours [6].

Our technology, which is based on machine learning, has an interface that highlights a few additional capabilities, such as Instagram and Facebook. The primary goal of the development was to create a user-friendly website for those who are reluctant to discuss their concerns and therapy-related issues after receiving a cancer diagnosis (at any stage), going through chemotherapy, or finishing treatment.

Even after successful operations and treatments, the healed group does not feel like the rest of us. Our goal is to break down this barrier and show that "The Cancer Beings" are just like regular humans and shouldn't be treated any differently.

II. PROBLEM IDENTIFICATION

In recent years, cancer has become a common topic of conversation among people. Daily updates are required for cancer therapy propaganda. The medical regimen and technology utilised to run the equipment used to find cancer cells must be developed in accordance with the many suggestions made. Our goal is to develop a web application that will enable access to influential and uplifting content for cancer sufferers while also assisting doctors in detecting cancer using our method.

Cancer research has seen a steady change throughout the past few decades. Researchers have used a variety of techniques, such as early-stage screening, to identify cancer types before symptoms appear. Additionally, they have created fresh methods for anticipating the results of cancer treatments early on. Large volumes of cancer data have been gathered and are available to the medical research community thanks to the development of new technologies in the field of medicine. However, one of the most fascinating and difficult problems for doctors is making an accurate prediction of a disease's course. As a result, ML techniques are now widely used by scientists conducting research in medicine. Many scientists only study one body organ at a time to look for cancer. However, in this research, we aim to create a system that can identify many cancer diseases in the human body using a single machine learning platform.

III. OBJECTIVE

1) This study project's primary goals are to identify the earliest stages of cancer diagnosis (such as those in the brain, breast, and lung) and investigate the degree to which machine learning algorithms are accurate.

2) Early identification of lung, brain, and breast cancer at its earliest stages.

It will also provide a suitable solution at an early stage and provide doctors with good software to diagnose cancer and their causes. Patients can also receive timely consultation.

IV. LITERATURE SURVEY

Over 1 crore people died last year, and breast cancer (BC) was to blame for 22.6% of those deaths. In India, BC accounts for 14.7% of all cancer cases and is the most prevalent malignancy in women. Numerous studies have been done on early BC detection, which can aid with timely treatment initiation and lower mortality rates. Only roughly 86% of cases are appropriately diagnosed out of all those that are. The risk of erroneous detection in cell biopsy pictures puts a person's life in danger. It is imperative to find new, alternative approaches that are simple to use with various data sets, economical, dependable, and secure, and capable of making correct predictions. Here, we offer a model that combines a number of Machine Learning (ML) algorithms, including Support Vector Machine, K-Nearest Neighbour, Decision Trees, and Artificial Neural Networks [5].

Cancer diagnosis and detection are becoming more and more automated. In the future, cancer prognosis will be relatively simple, and we won't even need to visit the hospital. As we can see, the medical industry is using and testing a variety of technologies. So, based on this, we can conclude that it will be simpler for us to identify cancer in the future. Among CART, SVM, and KNN, we are evaluating which algorithm will produce the best results. We are utilising machine learning to forecast three different types of cancer, including breast cancer, brain tumours, and lung cancer. FFor example, while predicting the type of breast cancer, we take into account characteristics such clump thickness, uniform cell size, uniform cell shape, etc[5][6].

The early detection and diagnosis of cancer disease has been made possible by the adoption of several medical imaging techniques. These photos can be utilised to detect cancer in its early stages, keep tabs on the situation, and follow up with cancer patients. It takes a lot of effort and time for doctors to interpret the various medical images. Medical picture interpretation performed manually is subject to prejudice and mistakes. The development of computer-aided diagnosis, which automates the interpretation of medical pictures for cancer early detection, began in 1980.By precisely and efficiently combining medical images, computer-aided diagnosis assists the doctor in identifying malignant disease in its early stages. The computer-aided diagnosis of cancer sickness has undergone numerous advancements in recent years.

Deep learning methods assist medical professionals in early cancer disease detection from medical imaging. For these algorithms to recognise comparable patterns in fresh medical images, machine learning methods need to be trained. Based on their learning technique, deep learning systems can be categorised into four groups: supervised learning, semi-supervised learning, unsupervised learning, and reinforced learning. Labelled data is necessary for the supervised learning method in order to train the machine learning algorithm. The label of unlabeled samples is additionally predicted using the trained model. Data without labels are necessary for the unsupervised learning method. To classify unknown samples, it builds an implicit model based on the training data. Both labelled and unlabelled data are used in the semi-supervised learning technique to train the algorithm. Through feedback, the reinforced learning approach learns from the environment. With practise, they become more effective [6][7][8].

G. Hemanth et.al. 2019 IEEE, The study suggests a method for automatic segmentation that uses CNNs (Convolution Neural Networks) to create tiny 3 x 3 kernels. Segmentation and classification are achieved with just this one technique. CNN (a machine learning technology) is based on layer-based neural networks for results classification. Data collection, pre-processing, average filtering, segmentation, feature extraction, and CNN via classification and identification are only a few of the stages that are involved in the proposed processes. Significant relations and patterns from the data can be extracted by using DM (data mining) techniques. Early brain tumour identification and prevention are being achieved with the use of data mining and machine learning approaches.

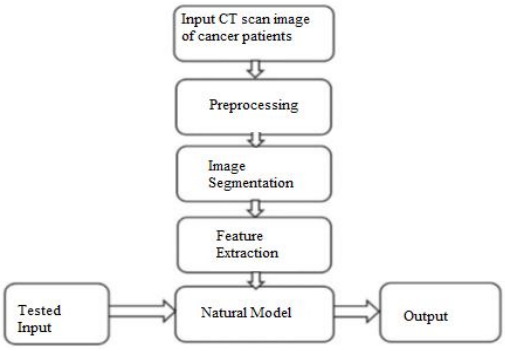
Meraj Begum Shaikh Ismail et. al. 2021, Turkish Journal, The main goals of this study are to identify the early stages of lung cancer and investigate the degree of accuracy of various machine learning algorithms. A thorough review of the literature revealed that certain classifiers have low accuracy while others have higher accuracy but are challenging to approach nearer to 100%. Due to inappropriate handling of DICOM pictures, implementation costs are expensive and accuracy is poor. Many various types of images are utilised in medical image processing, however CT scans are typically chosen since they have less noise. Deep learning has been shown to be the most effective technique for processing medical images, identifying and classifying lung nodules, extracting features, and predicting the stage of lung cancer. This system's initial stage employed image processing methods to separate lung sections. Using K Means, the segmentation is carried out. Segmented images are used to extract the features, which are then classified using a variety of machine learning algorithms. Based on their accuracy, sensitivity, specificity, and classification speed, the proposed techniques' performances are assessed.

Yash Amethiya et. al. 2021 , Elsevier, The goal of this review was to present several methods for looking at the use of various algorithms based on a machine learning (ML) approach and biosensors for the early diagnosis of breast cancer. Automation is required since ML and biosensors are required to recognise tumours from microscopic images. The goal of ML is to help computers learn for themselves. It is based on finding patterns in observed data and creating models to predict outcomes rather than relying on explicit pre-programmed rules and models.

The early detection and diagnosis of cancer disease has been made possible by the adoption of several medical imaging techniques. These photos can be utilised to detect cancer in its early stages, keep tabs on the situation, and follow up with cancer patients. It takes a lot of effort and time for doctors to interpret the various medical images. Medical picture interpretation performed manually is subject to prejudice and mistakes. Computer-aided diagnosis, which automates the interpretation of medical pictures for cancer early detection, has been in use since 1980. By precisely and efficiently combining medical images, computer-aided diagnosis assists the doctor in identifying malignant disease in its early stages.

V. METHODOLOGY

The system uses machine learning for its testing and training processes. The suggested model looks,



The block diagram in figure 4.1 above is categorised as follows:

A. *Pre-processing*

Pre-processing entails modifying the quality of the raw CT picture, as seen in fig. 1. This entails applying particular techniques to the image in order to enhance certain visual elements and data.

B. *Image Segmentation*

The term "image segmentation" refers to the division of a digital image into distinct parts. Pixels or super pixels are the same thing as photo segments. To portray an image more clearly, or in a way that is more significant and understandable, segmentation is utilised. A set of segments that jointly cover the full image are produced when an image is segmented, or a group of contours that are actually drawn from the image are produced. Every pixel in a region is equivalent to a characteristic or computed property, such as colour, intensity, or texture. Regarding the same property, the colour of neighbouring portions varies greatly. In medical imaging, which is common, the contours created following image segmentation can be used with a stack of images, to produce 3D reconstructions using interpolation methods like marching cubes.

C. *Feature Extraction*

The process of converting unprocessed numerical features from raw data while maintaining the integrity of the original data set's content. It yields superior results when compared to utilising machine learning on the raw data directly.

D. *Neural Model*

Neural networks are basic representations of how the nervous system works. An efficient illustration of how the human brain works is a neural network. It functions by simulating a vast number of interconnected processing units that resemble abstract models of neurons. Neural networks, a subset of machine learning that are often referred to as artificial neural networks (ANNs) or simulated neural networks (SNNs), are the foundation of deep learning approaches. Their organisation and nomenclature are modelled after the human brain, mimicking how organic neurons communicate.

VI. CONCLUSION

Recent research have shown good outcomes when using machine learning to diagnose cancer, with better generalization and accuracy than traditional methods. To address issues including the dearth of labelled data, selecting pertinent features, and model adaptation, more research is necessary.

In this review, we reviewed ML ideas and described how they are used in cancer prognosis and prediction. The majority of the studies that have been put forth in recent years concentrate on the creation of prediction models employing supervised machine learning techniques and classification algorithms with the goal of predicting accurate illness outcomes. It is clear from a review of their findings that combining multidimensional heterogeneous data with the use of various feature selection and classification approaches can result in useful tools for inference in the cancer domain.

VII. ACKNOWLEDGMENT

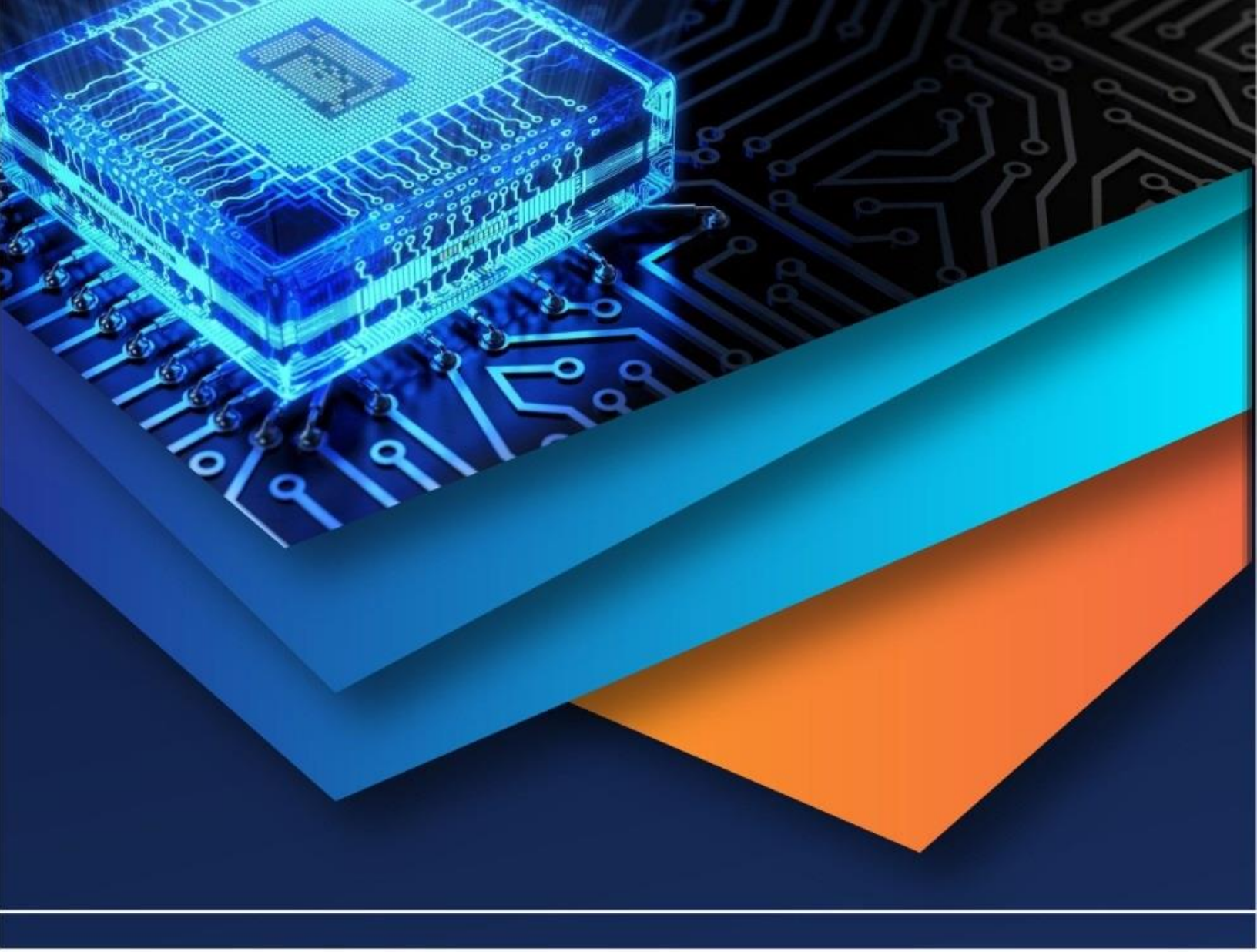
We take this opportunity to express our profound gratitude and deep regards to Our Project Guide , Department of Artificial Intelligence and machine learning, Tulsiramji Gaikwad Collage of Engineering & Technology, Nagpur, which provided guidance and space for us to complete this work.

REFERENCES

[1] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty and A.-B. M. Salem, “Classification using deep learning neural networks for brain tumors,” Future Computing and Informatics Journal, pp. 68-71, 2018.
[2] S. Bauer, C. May, D. Dionysiou, G. Stamatakos, P. Buchler and M. Reyes, “Multiscale modeling for Image Analysis of Brain Tumor Studies,” IEEE Transactions on Biomedical Engineering, vol. 59, no. 1, pp. 25-29, 2012.
[3] Islam, S. M. Reza and K. M. Iftekharuddin, “Multifractal texture estimation for detection and segmentation of brain tumors,” IEEE Transactions on Biomedical Engineering, pp. 3204-3215, 2013.
[4] M. Huang, W. Yang, Y. Wu, J. Jiang, W. Chen and Q. Feng, “Brain tumor segmentation based on local independent projection-based classification,” IEEE Transactions on Biomedical Engineering, pp. 2633-2645, 2014.
[5] Hamamci, N. Kucuk, K. Karaman, K. Engin and G. Unal, “Tumor-cut: Segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications,” IEEE Transactions on Medical Imaging, pp. 790-804, 2012.

[6] H. B. Menze, A. Jakab, S. Bauer, K. Farahani and K. Justin, "The Multimodal Brain Tumor Image Segmentation Benchmark," IEEE Transactions on Medical Imaging, pp. 1993-2024, 2014.
[7] Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu and Yi Pan, "A survey of MRI-based brain tumor segmentation methods," Tsinghua Science and Technology, pp. 578-595, 2014.
[8] S. Huda, J. Yearwood, H. F. Jelinek, M. M. Hassan, G. Fortino and M. Buckland, "A hybrid feature selection with ensemble classification for Imbalanced Healthcare Data: A case study for brain tumor diagnosis," IEEE Access, pp. 9145-9154, 2016.
[9] J. Seetha and S. S. Raja, "Brain tumor classification using Convolutional Neural Networks," Biomedical and Pharmacology Journal, pp. 1457-1461, 2018.
[10] P. Meena and V. Janani, "IMAGE SEGMENTATION FOR TUMOR DETECTION," A Monthly Journal of Computer Science and Information Technology, pp. 244-248, 2013.
[11] G.Hemanth et.al. 2019 IEEE , "Design And Implementing Brain Tumor Detection Using Machine Learning Approach".
[12] Meraj Begum Shaikh Ismail et. al. 2021, Turkish Journal of Computer and Mathematics Education ," Lung Cancer Detection and Classification using Machine Learning Algorithm".
[13] Yash Amethiya et. al. 2021 , Elsevier, "Comparative analysis of breast cancer detection using machine learning and biosensors".

.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY