TECH-CHRONICLE
AN INTERNATIONAL E-JOURNAL ON
EMERGING TRENDS IN SCIENCE, TECHNOLOGY AND MANAGEMENT

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

Bhushan S. Jakkanwar
PG Scholar
Artificail Intelliegence & Machine Learning
Tulsiramji Gaikwad Patil Collage of Engineering & Technology , Nagpur , India

Prof. Jayant Rohankar
Project Guide
Artificail Intelliegence & Machine Learning
Tulsiramji Gaikwad Patil Collage of Engineering & Technology , Nagpur , India

Prof. Snehal Bagde
Project Co-Guide
Artificail Intelliegence & Machine Learning
Tulsiramji Gaikwad Patil Collage of Engineering & Technology , Nagpur , India

Abstract- The illness known as cancer is said to include a wide variety of subtypes. The clinical patient treatment of cancer patients, which has become a prerequisite in cancer research, depends on early cancer identification and prognosis. Due to the need of categorising cancer patients into high or low risk groups, numerous research teams from the biomedical and bioinformatics sectors have looked into the application of machine learning (ML) approaches. These techniques have been used to simulate the beginning and treatment of cancerous disorders. The ability of ML algorithms to identify significant features in complicated datasets further demonstrates their importance. The field of cancer research has extensively used Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs), and Decision Trees (DTs) to build prediction models that allow precise and efficient decision-making. It is obvious that using ML approaches can improve our comprehension of how cancer spreads, but more study is required before these techniques can be utilised to direct ordinary clinical care. The detection of cancer in numerous physiological organs, including the brain, breast, and lung, is modelled in this work.

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

I. INTRODUCTION

The study of cancer has steadily advanced during the past few decades. Researchers have used a variety of techniques, such as early-stage screening, to detect various cancer types before symptoms appear. They have also created innovative methods for predicting the results of cancer therapy early on. The scientific community now has access to a lot of knowledge about cancer thanks to the development of modern medical technology [1]. To provide a precise forecast of the progression of a disease, however, is one of the most fascinating and difficult problems that doctors today are faced with. As a result, ML techniques are becoming popular among scientists working in the field of medicine. The trajectory of a particular tumour can be precisely predicted using these methods, and patterns and correlations can be gleaned from huge databases [1][2].

We give a review of studies using these methods in relation to the prediction and prognosis of cancer in light of the significance of personalised medicine and the emerging trend on the use of ML approaches. These studies consider prognostic and predictive factors that might not be connected to a particular drug or that might be combined to suggest a course of treatment for cancer patients, respectively. Additionally, we go over the numerous machine learning (ML) techniques being employed, the different kinds of data they use, the effectiveness of each suggested technique, as well as their advantages and disadvantages [3].

The utilisation of mixed data, including clinical and genetic data, is a definite trend in the research that have been suggested. But a recurring issue we identified in most works is the absence of external validation or testing on the veracity of their models. It is obvious that using ML techniques could increase the precision of survival, cancer susceptibility, and recurrence predictions. Utilising ML approaches, the accuracy of cancer outcome prediction has increased by 15%–20% over the last few years[4].

"Being healthy" is everyone's main priority. The right to a healthy life belongs to everyone. Our attitude is usually influenced by our feelings. everything, whether it's good or bad! It is essential to remember to keep a "+ve Approach" to health under these circumstances [5].

Cancer is one such illness that has a large impact on our society. It is usually a deadly syndrome brought on by a confluence of several clinical changes and genetic illnesses. Cancerous cells have the ability to proliferate and generate fatal growths everywhere in the human body. As a result, our project's main goals are to improve present practises and offer a platform for cancer diagnosis. Currently being developed is a web programme that will assist every cancer survivor. Users of this programme will gain understanding about common cancer types, their symptoms, dos and don'ts for treating each type, financial challenges, and other real-world difficulties. Motivational forum showcasing the latest recent ML and DIP techniques for spotting brain tumours and breast cancer [6].

The interface of our machine learning-based system highlights a few additional functionalities, such as Instagram and Facebook. The major goal of the project was to develop 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 a successful operation and subsequent therapies, the recovered 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 regular people who shouldn't be treated any differently.

II. PROBLEM IDENTIFICATION

The subject of cancer has become more often discussed in recent years. Updates on a daily basis must be included in cancer therapy propaganda. When creating a treatment plan and the technology that powers the tools used to detect cancer cells, the many recommendations made must be taken into account. In order to help medical professionals spot cancer using our technology, we want to develop a web application that gives cancer sufferers access to motivating and influencing information.

The study of cancer has steadily advanced during the past few decades. Researchers have used a variety of techniques, such as early-stage screening, to detect various cancer types before symptoms appear. They have also created innovative methods for predicting the results of cancer therapy early on. The scientific community today has access to vast amounts of information regarding cancer as a result of the development of modern medical technologies. To provide a precise forecast of the progression of a disease, however, is one of the most fascinating and difficult problems that doctors today are faced with. As a result, ML techniques are widely used by scientists conducting medical research. When hunting for cancer, many researchers only concentrate on one area of the body at a time. The objective of this research, on the other hand, is to create a system that can identify several types of cancer in the human body using a single machine learning platform.

III. OBJECTIVE

• The primary goals of this study are to assess the precision of machine learning algorithms and to recognise the earliest stages of cancer diagnosis (such as those in the breast, lung, and brain).
• The early identification of breast, lung, and brain cancer.
Additionally, it will provide a suitable treatment at an early stage and provide medical professionals cutting-edge software to recognise the causes of cancer and its symptoms. Patients can also receive timely consultation.

IV. LITERATURE SURVEY

22.6% of the nearly 1 crore deaths that occurred last year were due to breast cancer (BC). BC accounts for 14.7% of all cancer cases in India and is the most prevalent malignancy among women. Numerous studies have focused on early BC detection, which can aid in quick treatment initiation and lower mortality rates. Only 86% of cases are successfully diagnosed out of all those that are. Due to the potential for inaccurate detection in cell biopsy images, a person's life is in danger. It is vital to find new, alternative techniques that are simple to use with various data sets, inexpensive, trustworthy, and secure, and capable of producing precise forecasts [5].

Cancer diagnosis and detection are becoming more automated. We won't even need to visit the hospital in the future because the prognosis for cancer will be rather simple. As we can see, the medical industry is utilising and testing a variety of technologies. This suggests that cancer detection will likely get simpler in the future. CART, SVM, and KNN are the three algorithms whose performance we are evaluating. Machine learning is being used to forecast three types of cancer, including lung cancer, brain tumours, and breast cancer. As an example, when identifying the breast cancer kind, we take into account elements like clump thickness, uniform cell size, uniform cell shape, etc [5][6].

The widespread use of a variety of medical imaging tools has simplified the diagnosis and early cancer detection process. These images can be used for condition monitoring, cancer patient follow-up, and early cancer detection. Interpreting the multiple medical images requires a lot of work and time from doctors. The interpretation of medical images manually is subject to bias and mistakes. The development of computer-aided diagnostics, which automates the interpretation of diagnostic images for the early diagnosis of cancer, began in 1980.By accurately and successfully integrating medical images, computer-aided diagnostics enables the doctor to identify malignant disease in its early phases. Significant advancements in computer-aided cancer diagnosis have been made recently.

Using medical imaging and deep learning algorithms, medical professionals can identify early cancer sickness. These algorithms must be trained to recognise similar patterns in current medical images using machine learning techniques. Based on their learning processes, deep learning systems can be categorised into four groups: supervised learning, semi-supervised learning, unsupervised learning, and reinforced learning. Labelled data are needed to train the machine learning algorithm in the supervised learning approach. The label of unlabeled samples is likewise predicted using the trained model. The unsupervised learning method requires unlabeled data. In order to classify unidentified samples, it builds an implicit model based on the training data. Using both labelled and unlabeled data, the algorithm is taught using the semi-supervised learning technique. Through feedback, the reinforced learning technique gathers information from the environment. With practise, they become more proficient [6][7][8].

G. Hemanth et.al. 2019 IEEE, The study suggests a method for automatically segmenting images that creates tiny 3 x 3 kernels using CNNs. Segmentation and classification can both be done using a same approach. CNN (a machine learning technology) is built on layer-based neural networks, and it uses them to categorise results. Data collection, pre-processing, average filtering, segmentation, feature extraction, and CNN via classification and identification are only a few of the processes that the proposed methodologies cover. Techniques for data mining can be used to draw out important linkages and patterns

from the data. Early detection and prevention of brain tumours are being achieved using data mining and machine learning techniques.

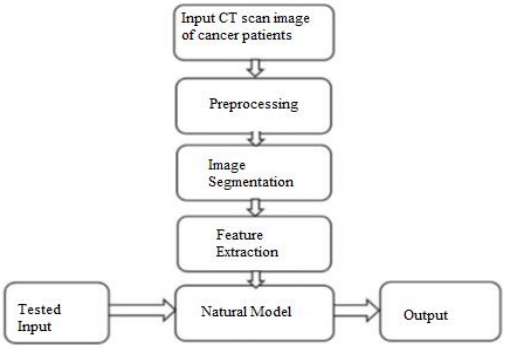
Meraj Begum Shaikh Ismail et. al. 2021, Turkish Journal, The main goals of this study are to identify the early stages of lung cancer and evaluate the precision of several machine learning methods. A careful review of the literature revealed that some classifiers are accurate but difficult to approach 100%, while others are more accurate. When DICOM images are processed incorrectly, huge implementation costs and poor accuracy ensue. A wide range of images are utilised in medical image processing, although CT scans are typically used 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 attributes, and figuring out the stage of lung cancer. K Means is used to complete the segmentation. Segmented images are used to extract the features, which are subsequently classified using a variety of machine learning methods. In terms of accuracy, sensitivity, specificity, and classification speed, the suggested strategies are assessed.

Yash Amethiya et. al. 2021 , Elsevier, The purpose of this review was to present different methods for investigating the use of different machine learning (ML)-based algorithms and biosensors for the early diagnosis of breast cancer. Automation is required since ML and biosensors are required to recognise tumours from microscopic images. ML aims to support the self-learning of computers. 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 widespread use of a variety of medical imaging tools has simplified the diagnosis and early cancer detection process. These images can be used for condition monitoring, cancer patient follow-up, and early cancer detection. Interpreting the multiple medical images requires a lot of work and time from doctors. The interpretation of medical images manually is subject to bias and mistakes. In order to automatically assess medical images for the early detection of cancer, computer-aided diagnosis has been employed since 1980. By precisely and successfully combining medical images, computer-aided diagnostics enables the doctor to detect malignant disease in its early phases.

V. METHODOLOGY

For its testing and training procedures, the system makes use of machine learning. The proposed model is discernible,



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

Pre-processing

As seen in fig. 1, pre-processing entails improving the quality of the raw CT image. This entails improving specific visual elements and information in the image using particular techniques.

Image Segmentation

"Image segmentation" is the process of dividing a digital image into discrete parts. Pixels and super pixels are two terms used to describe photo parts. Segmentation is a technique used to make an image more meaningful and understandable or to communicate it more effectively. When an image is segmented, a set of segments or a set of contours that are really drawn from the image are formed that collectively cover the full image. Each pixel in a region has the same feature, such as colour, intensity, or texture, that was calculated. The adjacent parts of the same piece of land are significantly dissimilar in colouring. Image segmentation contours, which are frequently employed in medical imaging, can be combined with a stack of photos to produce 3D reconstructions using interpolation techniques like marching cubes.

Feature Extraction

converting unprocessed numerical characteristics from raw data while maintaining the integrity of the content of the original data set. It yields better results in comparison to performing machine learning on the raw data directly.

Neural Model

Neural networks are fundamental representations of how the nervous system operates. A useful illustration of how the human brain works is a neural network. This is how it works: it mimics a massive number of connected processing units that resemble simplified models of neurons.

Neural networks, also referred to as artificial neural networks (ANNs) or simulated neural networks (SNNs), are the foundation of deep learning techniques. They are named and organised similarly to how organic neurons in the human brain communicate.

VI. CONCLUSION

Recent research has shown that machine learning outperforms traditional methods in terms of generality and

accuracy when used to identify cancer. To address issues such the absence of labelled data, selecting pertinent features, and model adaptation, more research is necessary.

In this review, we looked into ML ideas and analysed how they relate to cancer prognosis and prediction. The majority of studies that have been reported in recent years have concentrated on developing prediction models utilising supervised machine learning strategies and classification algorithms with the aim of properly forecasting illness outcomes. According to a review of their work, combining multidimensional heterogeneous data with various feature selection and classification methodologies could result in useful inference-making tools for the cancer domain.

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".

.