



A comprehensive analysis of recent advancements in cancer detection using machine learning and deep learning models for improved diagnostics

Hari Mohan Rai¹ · Joon Yoo¹

Received: 20 June 2023 / Accepted: 26 July 2023 / Published online: 4 August 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Purpose There are millions of people who lose their life due to several types of fatal diseases. Cancer is one of the most fatal diseases which may be due to obesity, alcohol consumption, infections, ultraviolet radiation, smoking, and unhealthy lifestyles. Cancer is abnormal and uncontrolled tissue growth inside the body which may be spread to other body parts other than where it has originated. Hence it is very much required to diagnose the cancer at an early stage to provide correct and timely treatment. Also, manual diagnosis and diagnostic error may cause of the death of many patients hence much research are going on for the automatic and accurate detection of cancer at early stage.

Methods In this paper, we have done the comparative analysis of the diagnosis and recent advancement for the detection of various cancer types using traditional machine learning (ML) and deep learning (DL) models. In this study, we have included four types of cancers, brain, lung, skin, and breast and their detection using ML and DL techniques. In extensive review we have included a total of 130 pieces of literature among which 56 are of ML-based and 74 are from DL-based cancer detection techniques. Only the peer reviewed research papers published in the recent 5-year span (2018–2023) have been included for the analysis based on the parameters, year of publication, feature utilized, best model, dataset/images utilized, and best accuracy. We have reviewed ML and DL-based techniques for cancer detection separately and included accuracy as the performance evaluation metrics to maintain the homogeneity while verifying the classifier efficiency.

Results Among all the reviewed literatures, DL techniques achieved the highest accuracy of 100%, while ML techniques achieved 99.89%. The lowest accuracy achieved using DL and ML approaches were 70% and 75.48%, respectively. The difference in accuracy between the highest and lowest performing models is about 28.8% for skin cancer detection. In addition, the key findings, and challenges for each type of cancer detection using ML and DL techniques have been presented. The comparative analysis between the best performing and worst performing models, along with overall key findings and challenges, has been provided for future research purposes. Although the analysis is based on accuracy as the performance metric and various parameters, the results demonstrate a significant scope for improvement in classification efficiency.

Conclusion The paper concludes that both ML and DL techniques hold promise in the early detection of various cancer types. However, the study identifies specific challenges that need to be addressed for the widespread implementation of these techniques in clinical settings. The presented results offer valuable guidance for future research in cancer detection, emphasizing the need for continued advancements in ML and DL-based approaches to improve diagnostic accuracy and ultimately save more lives.

Keywords Cancer detection · Machine learning · Deep learning · Feature extraction · State-of-art analysis · Brain tumor · Lung cancer · Skin cancer · Breast cancer · GLCM · CNN · SVM

✉ Hari Mohan Rai
drhmrai@gachon.ac.kr
Joon Yoo
joon.yoo@gachon.ac.kr

¹ School of Computing, Gachon University, 1342
Seongnam-daero, Sujeong-gu, Seongnam-si 13120,
Gyeonggi-do, Republic of Korea

Introduction

Cancer has been a persistent issue for humanity throughout recorded history. However, in recent decades, the incidence of cancer has significantly increased due to various factors. These factors include the accelerated aging of the population, the adoption of unhealthy habits by individuals, and the heightened presence of carcinogens in the environment, as well as in food and

beverages (Faguet 2015). The term “cancer” finds its earliest recorded use in the Greek language, where it was derived from the word “*kapkivoc*,” which meant both tumor and crab. The term “cancer” entered medical terminology in the 1600 s and refers to cells that have invaded or migrated to different parts of the body, potentially causing harm (Kumar et al. 2022). Given that the human body consists of trillions of cells, an infection can originate in any part of the body, ranging from the extremities to the brain. The process of cell division allows cells to multiply and fulfill the body's requirements. Normally, cells mature and die off naturally or due to damage, making way for new cells to take their place. However, in certain cases, this process malfunctions, leading to the proliferation or growth of damaged or abnormal cells when they should not be growing. These cells come together to form a mass of tissue called a tumor, which can be either cancerous (malignant) or noncancerous (benign) (Kumar et al. 2022). Understanding the nature of cancer, its causes, and the mechanisms behind abnormal cell growth is crucial for effective prevention, diagnosis, and treatment. Advances in research, including the exploration of artificial intelligence and machine learning techniques, have contributed to our understanding of cancer detection and diagnosis. By analyzing various types of cancer and the literature related to their detection and classification, researchers aim to improve the accuracy and effectiveness of cancer screening and treatment strategies.

Tumors can be classified into two broad categories: benign and malignant. Benign tumors are non-cancerous and are more common than malignant tumors. They do not have the ability to invade nearby tissues or spread to other parts of the body. Although benign tumors can sometimes grow to a significant size, they are generally not considered life-threatening, except in the case of brain tumors. Malignant tumors, on the other hand, have the potential to develop, infiltrate nearby tissues, and spread to distant parts of the body through a process called metastasis. This ability to metastasize is a characteristic feature of cancerous tumors. Most cancers manifest as solid tumors, but there are exceptions, such as blood cancers like leukemia, which do not form solid masses. Cancer is fundamentally a genetic disease, arising from abnormalities in genes that regulate cell growth and division. These genetic abnormalities disrupt the normal interactions and behaviors of cells. Normally, our body has mechanisms to eliminate cells with damaged DNA before they can become cancerous. However, as we age or due to other factors, these protective mechanisms may become less effective, increasing the risk of cancer development. Each individual's cancer has a unique pattern of genetic changes. As cancer progresses, additional genetic alterations occur, driving the growth and spread of the disease. It is possible for different cells within the same tumor to have their own genetic mutations, adding to the genetic complexity of the cancer. There are three main gene categories commonly influenced by genetic abnormalities in cancer: proto-oncogenes, tumor suppressor genes, and DNA repair genes. Proto-oncogenes are involved in the normal growth and division

of cells. However, when these genes undergo specific changes or become overactive, they can transform into oncogenes, which promote uncontrolled cell growth and contribute to the development of cancer. Tumor suppressor genes regulate cell proliferation and division, acting as “brakes” to prevent excessive cell growth. Mutations in these genes can disrupt their normal function, leading to uncontrolled cell division and an increased risk of cancer. DNA repair genes play a crucial role in maintaining the integrity of the DNA by fixing damage. Mutations in these genes impair the ability to repair DNA, increasing the likelihood of genetic abnormalities and subsequent mutations in other genes. These genetic alterations, including chromosomal abnormalities, contribute to the development of cancer. Metastasis occurs when cancer cells spread from the primary tumor site to other parts of the body. It signifies an advanced stage of cancer. Metastasis involves a complex process of cancer cells invading blood or lymphatic vessels, circulating through the body, and establishing new tumors in distant locations. This spread of cancer cells is a major challenge in cancer treatment and is associated with a poorer prognosis. Understanding the genetic basis of cancer and the specific gene alterations involved is crucial for developing targeted therapies and effective treatment approaches. Ongoing research continues to explore the genetic intricacies of cancer to improve diagnostics, treatment strategies, and ultimately, patient outcomes (NCI 2021).

Motivation

Cancer is a significant global health issue, contributing to a large number of deaths worldwide. Over the past decade, there has been a nearly 200 percent increase in cancer cases, making it one of the leading causes of mortality (Manhas et al. 2022). In 2019, cancer was the second-leading cause of death globally, accounting for 10.08 million deaths statistics (Burden and of Disease 2019). In the United States, cancer is also the second leading cause of death (Heron 2021; Jiaquan et al. 2019). The high mortality rate associated with cancer highlights the importance of early and accurate diagnosis, as well as effective treatment. The incidence of cancer has risen significantly over the years, particularly in developing and underdeveloped regions. Asia has a higher cancer mortality rate compared to other parts of the world (Ferlay et al. 2019). It is projected that by 2030, there will be 23.6 million new cancer cases diagnosed annually. The burden of cancer is disproportionately higher in less developed regions, including Central America, Africa, and parts of Asia. A majority of newly diagnosed cases and cancer-related fatalities occur in these regions. Several factors contribute to the rising incidence of cancer, including excess body fat, infections, exposure to ultraviolet light, alcohol consumption, and lifestyle choices. Access to medical treatment is a significant challenge, particularly for those living in poverty. Many individuals around the world do not have the means to afford medical care, including accurate diagnosis and appropriate treatment. Timely

diagnosis and access to affordable treatment can greatly improve survival rates and save lives. Developing advanced and cost-effective medical infrastructure is crucial in effectively managing and treating diseases like cancer. It is worth noting that the survival rate for specific types of cancer varies across countries. For example, the Republic of Korea has the highest survival rate for breast cancer, demonstrating the impact of a well-developed medical infrastructure on improving outcomes (GCO–SURV-CAN 2023). Addressing the global cancer burden requires concerted efforts in prevention, early detection, and accessible treatment options. It is essential to prioritize healthcare infrastructure development, increase awareness about risk factors, and ensure equitable access to healthcare services to combat the growing impact of cancer on individuals and societies worldwide.

The primary motivation for conducting this comprehensive assessment of cancer detection methods is the lethality of this illness, the alarmingly rapid rate at which cancer develops, and the alarmingly high mortality rate experienced by cancer patients worldwide (World Health Organization 2019; World Health Organization International Agency for Research on Cancer 2020). According to the World Cancer Report 2019 by the World Health Organization (WHO) and the International Agency for Research on Cancer (IARC), cancer is a major global health concern with increasing incidence and mortality rates (World Health Organization 2019). The Global Cancer Observatory (GCO), which is a database maintained by the WHO, provides valuable data and trends related to cancer incidence and mortality (World Health Organization 2019; World Health Organization International Agency for Research on Cancer 2020). Table 1 provides estimates for the number of new cancer cases and deaths from 2020 to 2040. The data shows a significant increase in both incidents and fatalities over the years. In 2020, there were approximately 19 million new cancer cases and 9.9 million deaths. By 2040, these numbers are projected to exceed 30 million cases and 16 million deaths. The risk percentages for incidents and fatalities are expected to reach 56.67% and 63.69%, respectively, by 2040. These statistics highlight the urgent need for preventive measures, early detection, and improved healthcare access to address the rising global cancer burden (World Health Organization 2019).

Figure 1 also from the GCO, provides a regional breakdown of cancer-related mortality rates for the year 2020. It shows that the Asia region is expected to have the highest

Mortality estimation in 2020 due to cancers

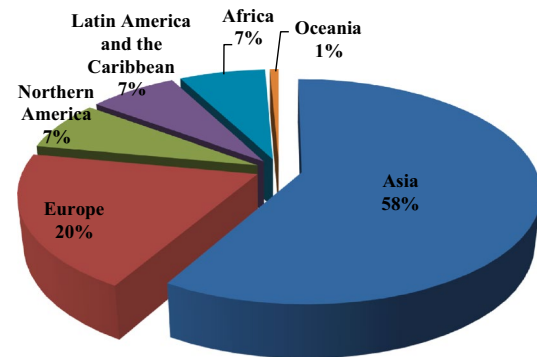


Fig. 1 The mortality rate estimation due to all cancer types, including both sexes, in 2020, continent-wise

death rate due to cancer at 58%, followed by Europe at 20%. The high death rate in Asia can be attributed to its large and dense population, accounting for approximately 59.4% of the world's total population. Additionally, dietary habits, such as alcohol consumption and smoking, contribute to the high cancer mortality rate in the region. In comparison, Europe has a lower population share (9.4%) but accounts for a higher proportion (20%) of cancer-related deaths. Africa, despite having a significant population (17.6% of the world's population), has a relatively lower percentage of deaths (7%) (United Nations Conference on Trade and Development 2022). These references support the notion that cancer is a global health concern with increasing incidence and mortality rates. The World Cancer Report 2019 and the Global Cancer Observatory provide comprehensive data on the global burden of cancer, highlighting the need for effective detection methods and interventions to address this growing public health challenge (World Health Organization 2019; The Global Cancer Observatory - All cancers 2020).

Above and beyond this, there are many patients who lose their lives as a result of manual or incorrect diagnostic errors. The most prevalent causes of medical litigation are diagnostic errors, which occur in 10–15% of cases on a regular basis and are responsible for 40–50 thousand fatalities each year (Harada et al. 2021). The empirical rates found in studies that make use of evaluations of charts are often one or more orders of a certain

Table 1 The estimation of Number of new cases of cancer and the fatality due to cancer with risk percentage from 2020 to 2040

Year	Prediction		Risk change		Risk percentage (%)	
	Incidences	Fatality	Incidences	Fatality	Incidences	Fatality
2020	19,292,789	9,958,133	0	0	0	0
2025	21,857,570	11,380,213	2,564,781	1,422,080	13.294	14.2806
2030	24,588,848	12,940,153	5,296,059	2,982,020	27.451	29.9456
2035	27,425,751	14,614,925	8,132,962	4,656,792	42.1555	46.7637
2040	30,226,151	16,300,829	10,933,362	6,342,696	56.6707	63.6936

magnitude smaller. For instance, the degree of error in a study conducted in hospitals of 7926 information gathered about hospitalizations and deaths was found to be 0.4%, whereas the error rate in an elementary care-based research of 1957 data was found to be 2.4%. Both studies were based on hospital data. The aforementioned figure represents over twelve million individuals in the United States whose primary care is negatively affected annually. In the event that similar proportions are seen in emergency departments and non-primary care ambulatory clinic visits, this would result in 31 million incorrect diagnoses being made each year (Newman-Toker et al. 2021). Just try to fathom the harm that would be caused if diagnostic mistakes were made throughout the world, not to mention the patients who would suffer both financially and physically as a result of these mistakes. Because of this, the incorrect diagnosis not only wastes time, money, and the patient's family's emotional and financial resources, but it also puts the patient's life in danger.

Consequently, it is necessary to develop a system for the automated early identification of cancer that is very accurate, quick, and dependable since it has the potential to save the lives of millions of people. The majority of the time, we are unable to diagnose cancer in its early stages, which increases the risk that it may progress to a state that is more life-threatening and ultimately result in death. It is very important to discover and diagnose potentially deadly illnesses as early as possible, particularly malignancies, since this allows for patients to get the appropriate therapy, which has the potential to save their lives. Because of improvements in early identification as well as treatment that is both inexpensive and effective, the mortality rate attributable to the second largest cause of death, cancer, fell by 1.9%, continuing a modest but continuous decreasing trend that has been ongoing since 1993 in USA (Jiaquan et al. 2021). Therefore, the solution to this issue and to limit the diagnostic error that often occurs as a result of manual diagnosis or devices that are prone to mistakes is to give a computer-aided diagnostics (CAD) system that must be very accurate, efficient, fast, and dependable. This is the only way to reduce the likelihood of diagnostic errors occurring and save many people's lives. Scientists, researchers, and medical professionals have been putting in a lot of effort over the past few decades in order to discover a solution for early diagnosis of life-threatening diseases like cancer. Although they have had some success in this endeavor, there is still a long way to go.

Related work

There are many researchers who have done extensive review for the detection of various types of cancers using machine learning and deep learning-based models. In this section, we have presented and done the comparative analysis of recently published review articles on various types of cancer detection techniques using machine learning and deep learning models.

Additionally, we have presented the findings and limitations of each literature in Table 2.

Al-shamasneh and Obaidallah (2017) conducted a literature review focusing on breast, lung, and brain cancer detection using machine learning (ML)-based classification methods. While they explored various performance parameters, its limitations included the evaluation of only three specific types of cancer, potentially lacking a comprehensive understanding of cancer detection in general. Moreover, they solely concentrated on ML-based classifiers, neglecting the deep learning (DL) based approach, which could have limited the range of techniques considered. Levine et al. (2019) performed a literature review covering skin, breast, lung, and multiple cancers with a focus on the radiology and diagnostic medicine workflow. They provided insights into building and deploying convolutional neural networks (CNN) for image analysis. However, they placed less emphasis on cancer detection itself and had a limited number of included literatures, potentially excluding essential studies on cancer detection. Huang et al. (2019) reviewed the literature related to various cancers, including prostate, breast, gastric, colorectal, solid tumor, and non-solid tumor cancers. They compared the prognosis and diagnostic performance of AI algorithms to human pathologists. Although they encompassed a wide range of cancer types, the amount of literature available for each cancer type was relatively small, affecting the depth of analysis. Additionally, they focused exclusively on AI algorithms and did not consider ML or DL classifiers, potentially limiting the scope of the analysis to a specific subset of approaches. Saba (2020) conducted a literature survey focusing on breast, brain, lung, liver, skin, and leukemia cancer detection. She explained benchmark datasets and evaluated various performance metrics, presenting structured flow diagrams for each type of cancer detection technology. However, the review did not separately study ML and DL-based classification, combining them under the umbrella of "classification based on ML and DL." This might limit insights into the specific strengths and weaknesses of each approach. Furthermore, there was no comparison between DL and ML-based approaches, which could have provided valuable comparative analysis.

Majumder and Sen (2021) utilized the applications of AI in oncopathology and translational oncology in their study and focusing on breast, lung, solid, and brain cancer. While providing insights into the broader aspects of AI utilization in cancer research, the review had limitations, including the limited number of cancer types covered, potentially not offering a comprehensive overview of cancer detection across various malignancies. Additionally, the small number of available literatures for each cancer type might have impacted the depth and breadth of analysis, and variations in performance parameter metrics among the literatures made direct comparisons challenging. Bin Tufail et al. (2021) presented the study covering astrocytic, breast, colorectal, ovarian, gastric, liver, thyroid, and multiple cancers, focusing on

Table 2 Comparative analysis of survey papers for the diagnosis and detection of various types of cancers

Year	References	Types of cancer	Findings	Limitations
[2017]	Al-shamasneh and Obaidallah (2017)	Breast, lung, brain	Enough literature included More performance parameters are explored ML classifiers are also described	Only three cancer kinds have been evaluated Only ML-based classification methods have been explored Deep learning models have not been compared to ML-based classifiers
[2019]	Levine et al. (2019)	Skin, breast, lung, multiple	Highlights of radiology and the diagnostic medicine workflow Outlines the process of building and deploying a CNN for image analysis	Less emphasis on cancer detection The survey contained a very small number of literatures
[2020]	Huang et al. (2020)	Prostate, breast, gastric, colorectal, solid tumor, non-solid tumor	They offered a prognosis and diagnostic performance comparison between AI algorithms and human pathologists Large number of cancer types	Only a few parameters have been evaluated Few literatures per cancer kind, maximum three ML and DL classifiers are not considered The literature reviewed is also less
[2020]	Saba (2020)	breast, brain, lung, liver, skin, leukemia	Benchmark datasets are also explained briefly All forms of performance metrics are evaluated A structure flow diagram of each type of cancer detection technology	Classification based on ML and DL has not been studied separately There has been no comparison of DL and ML-based approaches
[2021]	Majumder and Sen (2021)	Breast, lung, solid, brain	Along with cancer detection, they have demonstrated AI in oncopathology and translational oncology	There are few cancer kinds and literature Performance parameter metrics vary by literature
[2021]	Bin Tufail et al. (2021)	Astrocytic, breast, colorectal, ovarian, gastric, liver, thyroid, multiple	A separate description of publicly available datasets for cancer research Cancer detection and segmentation have also been discussed separately A good amount of performance metrics has been explored	Only DL-based cancer detection has been included Each cancer type has not been reviewed independently
[2022]	Kumar and Alqahtani (2022)	Breast cancer, brain cancer, lung cancer, skin cancer, prostate cancer, multiple	Discussed the many kinds of deep learning models and their architectures Additionally examined the literature depending on imagine type Also included a summary of the benchmark dataset	Only DL methods were taken into consideration for this study Different pieces of literature use different metrics for performance parameter
[2022]	Kumar et al. (2022)	Brain, breast, cervical, colorectal, kidney, liver, lung, oral, prostate, skin, stomach, thyroid	Covered the most variety of cancer types and used a large number of literatures Included all kinds of performance metrics Also presented the challenges of each literature	All cancer kinds are examined together; there is no separate review DL and ML approaches are also not taken into account separately

Table 2 (continued)

Year	References	Types of cancer	Findings	Limitations
[2022]	Painuli et al. (2022)	Breast, lung, liver, skin, brain, pancreatic	Detailed review of benchmark datasets for all cancer types Overall utilization shares of ML/DL approaches in all six cancer types of diagnosis Best ML/DL classifiers based on accuracy A large amount of literature	When reviewing the relevant literature, both the DL and ML techniques are examined together rather than individually
[2022]	Manhas et al. (2022)	Cervical, oral, breast, brain, cancer	A complete review of the literature on each kind of cancer The review has a greater number of references All performance parameters are covered Discussion of an extracted feature and dataset	DL and ML approaches are also not taken into account separately
[2022]	Ibrahim et al. (2022)	Lung, skin, brain, prostate, and breast cancer	A discussion of the various image datasets has been presented Various DL model architectures have also been demonstrated	Very little literature has been evaluated The research only included DL models
[2023]	Maurya et al. (2023)	Brain, cervical, breast, skin, lung	Included all forms of performance parameters Included the benefits and drawbacks of all the material evaluated Presented the challenges of each cancer type	The ML and DL techniques were not independently assessed and compared The dataset used was not described
[2023]	Mokoatle et al. (2023)	Lung, breast, prostate, colorectal	In addition to reviewing four kinds of cancer detection approaches, they proposed their own cancer detection techniques This work provided state-of-the-art sentence transformers (SBERT) and the unsupervised SimCSE approach for identifying cancer	Only four kinds of cancer are investigated The dataset study was not included There is just one type of assessment metric provided

publicly available datasets and discussing cancer detection and segmentation. While exploring a variety of performance metrics and providing insights into the evaluation of cancer detection methodologies, the review had limitations. The study solely focused on DL-based cancer detection, excluding ML-based approaches, and did not review each cancer type independently, potentially limiting the depth of analysis for specific malignancies. Kumar and Alqahtani (2022) conducted a literature review focusing on breast cancer, brain cancer, lung cancer, skin cancer, prostate cancer, and multiple cancers. They discussed various deep learning models and their architectures, depending on image types like mammograms and histopathology images. The review provided a comprehensive overview of DL approaches for cancer detection, but it had some limitations. It exclusively focused on DL methods, neglecting other techniques such as ML or traditional approaches. Additionally, different literatures used different metrics for performance parameters, making direct comparisons challenging and potentially impacting the overall assessment of DL models. Kumar et al. (2022) performed literature survey covering brain, breast, cervical, colorectal, kidney, liver, lung, oral, prostate, skin, stomach, and thyroid cancers, including a large number of literatures and exploring various performance metrics. The review provided a comprehensive overview of the literature available for each cancer type. However, it did not conduct separate reviews for each cancer type, potentially limiting the depth of analysis and the understanding of specific malignancies. Furthermore, the review did not differentiate between DL and ML approaches, which could have provided valuable insights into the strengths and weaknesses of each methodology.

Painuli et al. (2022) conducted a thorough literature review encompassing breast, lung, liver, skin, brain, and pancreatic cancers. The study provided a detailed analysis of benchmark datasets for each cancer type and explored the utilization of both machine learning (ML) and deep learning (DL) approaches. It identified the best classifiers based on accuracy, offering a comprehensive overview of ML/DL applications in cancer detection. However, the review evaluated DL and ML techniques together, potentially limiting insights into the specific strengths and weaknesses of each methodology. Manhas et al. (2022) performed a comprehensive literature review covering cervical, oral, breast, brain, and other types of cancer. The study included a greater number of references and provided a complete analysis of the literature on each type of cancer. It covered all performance parameters and discussed extracted features and datasets. However, similar to previous reviews, the distinction between DL and ML approaches was not considered separately, potentially limiting the analysis of specific methodologies. Ibrahim et al. (2022) focused on a literature review that centered on lung, skin, brain, prostate, and breast cancer. The study extensively discussed various

image datasets and demonstrated various DL model architectures. However, it evaluated a relatively small number of literatures, which may limit the overall comprehensiveness of the analysis. Additionally, the review exclusively focused on DL models, excluding other methodologies such as ML or traditional approaches. Maurya et al. (2023) conducted an in-depth literature review encompassing brain, cervical, breast, skin, and lung cancers. The study included all forms of performance parameters and presented the benefits and drawbacks of the evaluated material. It also discussed the challenges associated with each cancer type. However, the review did not independently assess and compare ML and DL techniques, limiting insights into their relative strengths and weaknesses. Furthermore, the dataset used in the review was not fully described, potentially affecting the reproducibility of the findings. Mokoatle et al. (2023) focused on a literature review centered on lung, breast, prostate, and colorectal cancers. The study not only reviewed four types of cancer detection approaches but also proposed its own cancer detection techniques using sentence transformers (SBERT) and the unsupervised SimCSE approach. However, the review had limitations, including a limited number of evaluated literatures and the lack of extensive validation or comparison to other state-of-the-art methodologies, which may limit the understanding of the proposed techniques' effectiveness.

Contributions

In this study, we conduct an in-depth and comprehensive review of the current and future prospects in the diagnosis and detection of various types of cancers using traditional machine learning and deep learning models. Our research focuses on analyzing the advancements and applications of these models in the field of cancer detection. The main contributions of this review article are summarized as follows:

Reviewed literatures: The review analyzed 130 research articles published between 2018 and 2023, focusing on the diagnosis and detection of various types of cancers. Among them, 56 articles employed ML techniques, while the remaining 74 utilized DL approaches.

Comparative analysis of review articles: In addition to individual research articles, the review compared findings and limitations from recently published review articles on cancer detection spanning from 2017 to 2023, offering a broader context for the state-of-the-art in cancer detection research.

Separate analysis of DL and ML techniques: The review conducted a separate analysis of research articles utilizing DL and ML techniques for cancer detection. By examining both paradigms, the study provided a comprehensive overview of the strengths and limitations of each approach in cancer detection.

State-of-the-Art analysis: A separate state-of-the-art analysis was conducted for the detection of four specific types of cancers: brain, lung, skin, and breast. Each cancer type was thoroughly examined in terms of techniques, models, and achieved detection accuracy.

Parameters in research articles: The review carefully scrutinized various parameters within the research articles, including authors, publication years, feature extraction methods, datasets used for detection, best classifiers employed, and corresponding accuracy levels achieved by the models.

Descriptive analysis of benchmark datasets: The review presented a descriptive analysis of benchmarked datasets specific to the four types of cancer discussed in the study. Accessible links to each dataset were provided for the convenience of researchers.

Formulative study of common types of preprocessing techniques: A formulative study of common types of preprocessing techniques was performed. The review outlined various data preprocessing methods employed in cancer detection, such as data normalization, image resizing, and feature scaling. This analysis shed light on the critical role of preprocessing in enhancing model performance.

Commonly utilized features for cancer detection: An overview of commonly used features for cancer detection was presented, accompanied by concise explanations of their relevance and significance in the field.

Key findings and challenges: The review delved into key findings and challenges associated with both ML and DL approaches for each type of cancer. This analysis provided a comprehensive understanding of the strengths and limitations of different techniques in cancer detection.

Comparative analysis of model performance: For each type of cancer, the study conducted a comparative analysis of the best performing models and worst performing models based on their detection accuracy. This comparison offered insights into the effectiveness of different approaches and models in cancer detection.

Best vs. worst techniques with % improvement in accuracy: The study conducted a thorough comparative analysis of the best and worst performing techniques for each type of cancer, based on their detection accuracy. Moreover, the percentage improvement in accuracy achieved by the best-performing techniques over the worst-performing ones was quantified. This analysis highlighted the impact of utilizing superior techniques in cancer detection.

Overall, this review article provides a comprehensive overview of the current state of research in cancer detection, highlighting the progress made using both traditional machine learning and deep learning models. The findings and analysis presented can serve as valuable insights for researchers and practitioners in the field, helping to advance the development of more accurate and efficient methods for diagnosing and detecting various types of cancers.

Research methodology

In summary, the process of detecting different forms of malignancies using machine learning and deep learning algorithms involves several steps. Figure 2 illustrates the architectural process flow diagram for cancer detection. In this study, four types of cancer (brain, lung, skin, and breast) were considered. Two main applications of artificial intelligence (AI) for cancer diagnosis were discussed: traditional machine learning methods and deep learning methods. While there are some common procedures in both approaches, each has its own specific techniques and advantages. For traditional machine learning algorithms, the key steps include dataset collection, dataset preprocessing, feature engineering, data preparation, and classification. Dataset collection is the initial requirement, and preprocessing techniques such as filtering, resizing, normalization, and image enhancement are applied to the collected data. Feature engineering involves selecting, reducing, fusing, or enhancing features to improve classification performance. Various machine learning classifiers are then used for binary or multi-class classification of cancer types. In deep learning-based detection, most of the processes are similar to traditional machine learning except for the feature engineering phase. However, some researchers also incorporate feature engineering techniques to enhance detection accuracy in deep learning models. In the subsequent sections, each phase of cancer detection using both machine learning and deep learning classifiers is explored in more detail, including their respective advantages and disadvantages.

Datasets

The acquisition of data is a prerequisite for any machine learning project, and various datasets may provide variable outcomes for classifiers and detectors depending on the input they are given. In a similar line, various datasets need unique approaches in order to identify and categorize diseases such as cancer. There is a wide variety of processing accessible for many sorts of data, including pictures, text, and signals, amongst others. It is essential to keep in mind that although it is possible to diagnose cancer using real-time as well as offline information, the former option is far more prevalent. In this section, we will present a high-level review of the most prominent benchmarked datasets for the detection of cancers, brain, lung, skin, and breast, in the general population at large. Concise explanations and descriptions of each dataset, including the overall data size and the modality used, have been incorporated. Table 3 provides detailed information about

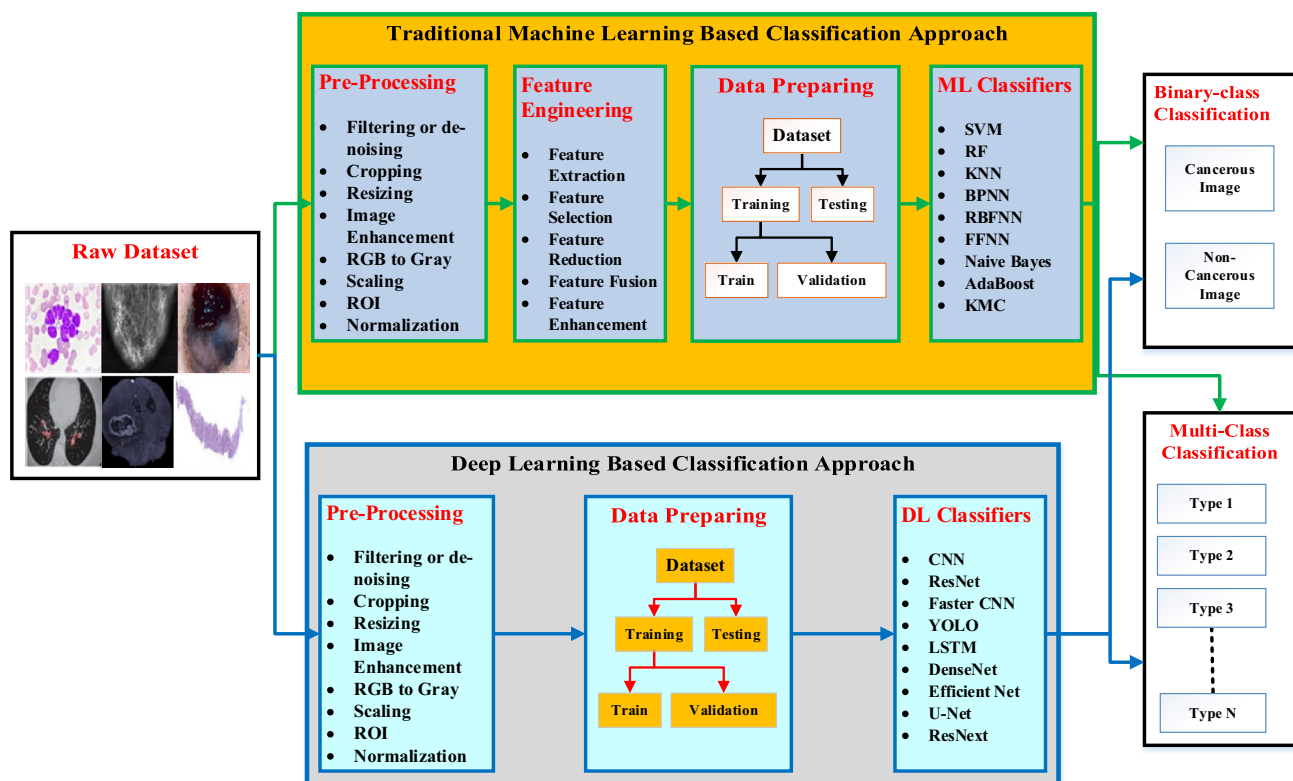


Fig. 2 The basic steps involved in the detection of cancer using ML and DL techniques

each dataset along with download links, allowing scholars to access and utilize them for their research purposes.

Brain tumor datasets

Researchers have access to various brain tumor datasets, including the widely used Brain Tumor Segmentation (BraTS) dataset. BraTS is an annual challenge that provides a benchmark for brain tumor identification and segmentation using multimodal MRI images. The dataset has evolved over the years, with BraTS 2021 containing 8000 mpMRI scans from 2000 patients (Bakas 2017; Menze et al. 2015; Baid et al. 2021). Other benchmark datasets, such as Figshare and The Cancer Genome Atlas of Glioblastoma Multiforme (TCGA-GBM), are also available. Figshare dataset includes 3064 T1-weighted MRI scans from 233 patients, categorized into different tumor types (Anaya-Isaza et al. 2023). The TCGA-GBM dataset comprises a large number of images (481,518) from 262 individuals and includes scans from various modalities (Scarpace 2016). These datasets enable researchers to investigate brain tumor identification and segmentation using different imaging modalities, contributing to advancements in diagnosis and treatment.

Lung cancer datasets

Several datasets are available for lung cancer research. The LIDC-IDRI dataset is widely used and contains 244,527 CT images from 1010 lung cancer patients (Armato et al. 2011). The LUNA-16 dataset is a subset of LIDC-IDRI and focuses on lung nodule identification with 888 CT scans (Jacobs et al. 2014). The Lungs Data Science Bowl Dataset (LDSBD-2017) dataset has 2101 CT scans for chest cavity analysis (Hamilton and “Data Science Bowl 2017”). The Iraq-Oncology Teaching Hospital-National Center for Cancer Diseases dataset, often known as the (IQ-OTHNCCD) dataset includes 1,190 CT images from 110 patients, categorized as benign, malignant, and normal cases (Alyasriy and AL-Huseiny 2021). These datasets offer valuable resources for studying lung cancer detection and advancing diagnostic methods.

Breast cancer datasets

The five widely used breast cancer datasets offer a diverse range of cases and features for research and analysis. The Wisconsin Diagnostic Breast Cancer (WDBC) (Mohammad et al. 2022) and Wisconsin Original Breast Cancer (OWBC) (Birchha and Nigam 2023) datasets provide case-based

Table 3 Description of benchmarked dataset for the detection of various cancer types

Dataset	Types of cancer	Download link	Description
BraTS	Brain	http://braintumorsegmentation.org/	It contains multi-parametric magnetic resonance imaging (mpMRI). Previously used versions are BraTS 2012, BraTS 2013, BraTS 2014, BraTS 2015, BraTS 2016, BraTS 2017, BraTS 2018, BraTS 2019, BraTS 2020, BraTS 2021. Its latest version BraTS 2021 contains 8000 mpMRI scans from 2000 subjects publicly available to use (Bakas 2017; Menze et al. 2015; Baid et al. 2021)
Figshare	Brain	https://figshare.com/articles/dataset/datasets/5472970	233 individuals with labeled images for pituitary tumor (930), meningioma (708), and glioma (1,426) on T1-weighted contrast-enhanced MRI are included in this dataset (Anaya-Isaza et al. 2023)
TCGA-GBM	Brain	https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=1966258	The Cancer Imaging Archive (TCIA) team has released the TCGA-GBM public dataset, which consists of 481,518 images obtained from CT and Digital Radiography (DX) modalities. These images are from 262 patients diagnosed with brain cancer (Scarpace 2016)
LIDC-IDRI	Lung	https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=1966254	The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) is a free and publicly available database with 244,527 annotated CT scans from 1010 people who were screened for lung cancer (Armato, et al. 2011)
LUNA-16	Lung	https://luna16.grand-challenge.org/Data/	The Lung Nodule Analysis 2016 (LUNA-16) is an open challenge for lung cancer detection and segmentation. This dataset is a subset of LIDC-IDRI benchmarked dataset, and this also includes CT images of lung nodules. In this dataset 888 CT scan of lung is available excluding the lung nodules of thickness greater than 2.5 mm (Jacobs et al. Feb. 2014)
LDSD-2017	Lung	https://www.kaggle.com/c/data-science-bowl-2017	The Booz Allen and Kaggle team published this dataset in two parts, the first including 1595 axial CT scans of lung cancer patients and the second 506 images. The National Cancer Institute (NCI) has made available a dataset consisting of CT scans of patients for use in a public research competition (Hamilton and “Data Science Bowl 2017)
IQ-OTHNCCD	Lung	https://data.mendeley.com/datasets/bhmdr45bh2/2	There are 1190 images in the collection, each one a CT scan slice from one of the 110 patients. Lung cancer patients at various stages of the disease and healthy volunteers are represented in this collection of CT images. Markings from oncologists and radiologists at these two institutions may be seen on IQ-OTHNCCD slides (Alyasriy and AL-Huseiny 2021)
WDBC	Breast	https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic)	There are a total of 569 unique cases and 32 corresponding characteristics in this data collection. There are 357 noncancerous instances and 212 cancerous ones (Mohammad et al. 2022)
OWBC	Breast	http://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(original)	There are a total of 699 cases in the Wisconsin Original Breast Cancer (OWBC) Dataset (65.52% of which are benign and 34.48% of which are malignant) (Birchha and Nigam 2023)
MIAS	Breast	https://www.repository.cam.ac.uk/handle/1810/250394	Using the MIAS database contains the original 322 images (161 pairs) captured at 50 micron resolution in "Portable Gray Map" (PGM) format, together with the corresponding truth data (Suckling et al. 2015)

Table 3 (continued)

Dataset	Types of cancer	Download link	Description
BCCD	Breast	https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra	The Breast Cancer Coimbra DataSet (BCCD) dataset includes 116 samples with 2 classes (64 patients and 52 healthy controls) and 9 attributes (i.e., glucose, resistin and insulin). All the attributes in Breast Cancer Coimbra data set are anthropometric data and parameters (Geetharamani and Sivagami 2021)
INbreast	Breast	https://www.kaggle.com/datasets/martholi/inbreast?select=inbreast.tgz	The INbreast dataset is comprised of a total of 410 mammographic images of breasts, which include both benign and malignant tumors. It includes 115 instances, 90 of which are from women who have both breasts (left and right), and each of those cases has four images. The other 25 cases are from patients who have had a mastectomy, and each of those cases contains just two images (Moreira et al. 2012)
ISIC	Skin	https://challenge.isic-archive.com/data/#2016	The 44,118 dermoscopy images were collected from 2746 patients and were divided for the ISIC 2020 dataset into a "train set" consisting of 33,136 images obtained from 2056 subjects and a "test set" consisting of 10,982 images obtained from 690 individuals (Cassidy et al. 2022)
Dermnet	Skin	https://www.kaggle.com/datasets/shubhamgoel27/dermnet	The Dermnet Dataset contains images depicting 23 distinct skin conditions; the total number of images in the dataset is around 19,500 (Bajwa, et al. 2020)
PH ²	Skin	https://www.fc.up.pt/addi/ph2%20database.html	This image collection contains 200 dermoscopy images of melanocytic lesions, including 80 images of common nevi, 80 images of atypical nevi, and 40 images of melanomas (Mendonca et al. 2013)
HAM10000	Skin	https://challenge.isic-archive.com/data/#2018	This dataset contains 10,015 dermatoscopic images from various populations and modalities and may be used as a training set for academic machine learning research. This belongs to the ISIC dataset (Tschandl et al. 2018a, 2018b)

clinical information without images, including details on benign and malignant cases. The Mammographic Image Analysis Society (MIAS) dataset consists of 322 breast images, with corresponding ground truth data and categorization into normal, benign, and malignant classes (Suckling et al. 2015). The INbreast dataset includes 410 mammograms, showcasing both benign and malignant cancers, with variations such as cases with two breasts, mastectomy patients, and corresponding images (Moreira et al. 2012). The Breast Cancer Coimbra DataSet (BCCD) is a dataset consisting of 116 samples with two classes: 64 breast cancer patients and 52 healthy controls. It includes nine attributes, primarily anthropometric data and parameters such as glucose, resistin, and insulin (Geetharamani and Sivagami 2021).

Skin cancer datasets

Four commonly used benchmarked datasets for skin cancer research are the ISIC dataset, Dermnet dataset, PH2 dataset,

and HAM10000 dataset. The International skin imaging cooperation (ISIC) dataset contains 44,118 dermoscopy images from 2746 individuals, with a separate training and test set (Cassidy et al. 2022). The Dermnet dataset offers 23,000 skin photos with labels from the Dermnet Skin Disease Atlas, although some images may be unavailable (Bajwa et al. 2020). The PH2 dataset includes 200 images of melanocytic lesions, categorized as melanomas, atypical nevi, and common nevi (Mendonca et al. 2013). Lastly, the Human Against Machine with 10,000 training images (HAM10000) dataset provides 10,015 dermatoscopic images for training machine learning models, featuring diverse demographics and modalities (Tschandl et al. 2018a, 2018b). These datasets serve as valuable resources for skin cancer detection and diagnosis studies.

Preprocessing

Pre-processing is a crucial step in image processing, object identification, and machine learning classification

problems. It involves transforming the raw data into a format that can be easily analyzed and interpreted by algorithms. Pre-processing helps to clean and normalize the data, removing any unwanted elements and filling in any missing values. This ensures that the data is accurate, complete, and consistent. Another reason why pre-processing is necessary is that it helps to reduce the dimensionality of the data, making it easier for the machine learning algorithms to process. This is particularly important when dealing with large datasets, as it can significantly reduce the time and computational resources required to train the models. Finally, pre-processing can help to extract relevant features from the data, which can be used for training machine learning models. This involves selecting the most important features and removing any redundant or irrelevant ones. By doing this, the models can be trained on the most relevant and informative data, leading to better performance and accuracy. The choice of pre-processing operations will depend on the specific problem and the characteristics of the data. It involves transforming raw data into a format that can be easily analyzed and interpreted by the algorithms, and it is essential for ensuring accuracy, reducing dimensionality, and extracting relevant features.

There are several types of pre-processing operations that can be performed on the data, depending on the type of data and the analysis that needs to be performed. Some common pre-processing operations include data cleaning, data normalization, data transformation, feature selection, dimensionality reduction, and many more.

Data study and analysis: This step involves exploring and analyzing the dataset to gain insights into its characteristics, such as the number of samples, features, and their distributions. It helps in understanding the dataset's properties and assists in making informed decisions about preprocessing techniques to be applied. No specific formula is associated with this step.

Data cleaning: Data cleaning involves handling noisy or erroneous data. Some common techniques include:

- **Removing duplicate entries:** Identify and remove duplicate rows from the dataset.
- **Dealing with missing values:** Fill in missing values using imputation techniques like mean, median, or mode imputation, or remove samples or features with missing values.
- **Correcting data errors or inconsistencies:** Identify and correct errors or inconsistencies in the dataset, such as formatting issues or incorrect values.

Data filtering: Filtering operations are applied to remove noise or artifacts from the dataset. Common techniques include:

- **Low-pass filtering:** A technique used to suppress high-frequency components in the data. It allows only the low-frequency components to pass through.
- **High-pass filtering:** A technique used to enhance high-frequency components in the data. It allows only the high-frequency components to pass through.

Filtering can be performed using various mathematical operations or filters such as the Fourier Transform, Butterworth filter, or Gaussian filter, depending on the specific requirements of the data.

Data transformation: Data transformation involves converting the original data into a different representation or form. Some common techniques include:

- **Image data conversion:** Convert RGB (Red, Green, Blue) images to grayscale using formulas like luminosity or average-based conversion.
- **Time series data transformation:** Apply mathematical operations like differencing or logarithmic transformation to make the data stationary or improve its properties.

Data normalization: Normalization scales the data to a standard range or distribution. Some common normalization techniques include:

- **Min-max normalization:** It is a widely used data preprocessing technique that scales the data to a specified range, often between 0 and 1. The formula for min-max normalization is given in Eq. (1). By applying this formula, the data is transformed into a normalized range, preserving the relative relationships between different data points while making it suitable for machine learning algorithms that benefit from input data being within a consistent range.

$$X_{\text{normalized}} = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})} \quad (1)$$

- **Clipping normalization:** Clipping normalization restricts the data within predefined minimum and maximum values, effectively bounding the data to a specific range. The generalized formula of clipping normalization is given in Eq. (2).

$$X_{\text{Cnormalized}} = \begin{cases} X_{\max}, & x > X_{\max} \\ X_{\min}, & x < X_{\min} \\ X, & \text{Otherwise} \end{cases} \quad (2)$$

- **Standard deviation normalization:** Standard Deviation normalization scales data by dividing each value by the standard deviation, ensuring that the data have a simi-

lar spread around the mean. The formula for standard deviation normalization is given in Eq. (3).

$$X_{SDnormalized} = \frac{X}{X_{std}} \quad (3)$$

- **Z-score normalization:** Z-score normalization, also known as standardization, is a data preprocessing technique that transforms the data to have zero mean and unit variance. It is achieved by subtracting the mean (X_{mean}) from each data point (X) and then dividing the result by the standard deviation (X_{std}) of the data. The formula for Z-score normalization is given in Eq. (4). This process ensures that the data is on a comparable scale and follows a standardized distribution, making it suitable for various machine learning algorithms that benefit from normalized data. By applying Z-score normalization, outliers are less likely to significantly influence the data, leading to improved model performance and more reliable results.

$$X_{standardized} = \frac{(X - X_{mean})}{X_{std}} \quad (4)$$

Data standardization: Data standardization is a crucial preprocessing technique used to transform data into a standard format. It aims to give the data a zero mean and a unit variance, making it easier for machine learning algorithms to process. The formula for standardization is the same as z-score normalization mentioned in Eq. (4). By applying this formula, the data is normalized to have a Gaussian distribution with a mean of 0 and a standard deviation of 1, facilitating more effective analysis and modeling.

Data resizing: Data resizing is a preprocessing technique used in image data analysis. It adjusts the size of images without altering their aspect ratio, ensuring compatibility with machine learning models or image analysis techniques. No specific formula is associated with resizing, as it depends on the desired target dimensions. The process maintains the image's content and appearance while standardizing its size for efficient processing and improved analysis performance.

Data cropping: Data cropping is a preprocessing technique used in image processing and computer vision tasks. It involves selecting a specific region of interest from an image and discarding the remaining unwanted portions. Unlike other normalization techniques, data cropping does not have a specific formula associated with it, as it depends on the desired region of interest and the specific task at hand.

Data scaling: Data scaling is a preprocessing technique used to transform data to a specific range or interval,

making it easier to compare and analyze data points with different scales. One common method for data scaling is linear scaling, which maps the data to a desired range. The formula for linear scaling is given by Eq. (5).

$$X_{scaled} = \frac{(X - X_{min})}{(X_{max} - X_{min})} * (new_max - new_min) + new_min \quad (5)$$

Data sampling: Data sampling is a crucial preprocessing technique used in various data analysis and machine learning tasks. It involves selecting a subset of data from the original dataset to create a representative sample for analysis. Different sampling strategies can be employed, such as random sampling, stratified sampling, or oversampling/undersampling for imbalanced datasets. There is no specific formula associated with data sampling, as the sampling process depends on the chosen strategy.

These preprocessing techniques help in preparing the data for further analysis by addressing noise, missing values, inconsistencies, or data-specific challenges. The choice of preprocessing techniques depends on the nature of the data and the objectives of the analysis or machine learning task at hand.

Feature engineering

Feature engineering is another very important and crucial step for classification problems, especially for traditional machine learning algorithms. Features are nothing but characteristics of the dataset based on which machine classifies the given task. Instead of complete dataset only features must be fed into the classifiers based on which the models used to provides the classified outcomes. As the feature engineering is performed on the pre-processed dataset, it must be assured that the data which is to be applied for extracting the features should be cleaned and corrected. But finding out what feature should be chosen to extract, which methods to be applied, in what domain it must be applied, how many features must be extracted, these are the most basic questions that come before applying feature engineering to the dataset. Also, the machine learning classifier performance can be improved upon the correct and better feature but in other side it may give worst performance in case of wrong and inaccurate features. The feature engineering can further be classified into several classes such as feature extraction, feature selection, feature reduction, feature fusion, and feature enhancement.

Feature extraction means extracting useful and meaningful information from the pre-processed dataset and these extracted features only fed to the classifier. Feature extraction can be handcrafted or manual or handmade or automated or computerized. The performing feature extraction task is again a tedious task especially when the dataset is

huge. The number features to be extracted also play an important role because very large features may confuse the machine learning model and too small features may not be able to correctly classify the data in an efficient way. Therefore, many researchers apply feature selection techniques, where among the entire extracted feature appropriate feature to be selected. For selecting the appropriate and correct feature again can have variety of techniques manual or automated. Similarly, feature reduction is another kind of feature engineering which plays a similar role as feature selection. It means extracted feature may be of huge number or it may not be normalized or scaled properly, that can be done using feature reduction technique. The mostly and majorly used feature reduction technique is principal component analysis (PCA). Furthermore, many researchers are working on feature fusion to improve the efficiency of classifiers in detection tasks. In case of feature fusion, if the feature is extracted, selected or reduced by multiple methods then it can be combined in same parameters. Feature enhancement also sometimes utilized such as optimization methods are one of the examples of the feature enhancement technique. There are similarly many methods by which feature enhancement can be performed more appropriately and efficiently. But it must be noted that feature engineering may not be compulsory where deep learning models are applied for the detection task because the DL models are constructed in such a way that they can automatically extract the feature from the provided dataset. Table 4 presents a comprehensive summary of the most commonly and prominently used features for cancer detection tasks, accompanied by their respective descriptions.

Classifiers

Machine learning models

The advancement of classification tools, particularly those based on machine learning (ML) and deep learning (DL), has significantly improved the ability to classify and diagnose a wide range of malignancies and disorders. ML models, including neural networks and traditional ML algorithms, require extracted information as input for effective performance. When it comes to cancer detection, ML-based classifiers have demonstrated high accuracy, especially when combined with non-invasive diagnostic techniques, making them one of the most effective approaches. While there have been many studies utilizing machine learning classifiers for cancer detection, the processes for data preparation and feature extraction vary across research studies and are not standardized. However, these classifiers have been successfully applied to identify different types of malignancies. Some of the commonly used ML classifiers for cancer prediction include:

- (1) *Support vector machine (SVM)*: Support vector machine is a supervised learning algorithm used for both classification and regression tasks. SVM finds an optimal hyperplane that separates data points belonging to different classes in a high-dimensional feature space. It aims to maximize the margin between the hyperplane and the nearest data points of each class. SVM can handle linear and nonlinear classification problems using different kernel functions, such as linear, polynomial, and radial basis kernel. In cancer detection, SVM has been widely used due to its ability to handle high-dimensional data and its flexibility in incorporating different kernel functions to capture complex relationships between features and target labels. By learning from labeled training data, SVM can build a decision boundary that can accurately classify cancerous and non-cancerous samples.
- (2) *Random forest (RF)*: Random forest is an ensemble learning algorithm that combines multiple decision trees to make predictions. Each decision tree in the random forest is constructed using a random subset of features and a random subset of training data. During training, the decision trees are built independently, and the final prediction is made by aggregating the predictions of all the trees through voting or averaging. Random Forest is known for its robustness and ability to handle high-dimensional data. It can capture complex relationships between features and target labels and handle noisy and missing data effectively. Random Forest also provides feature importance measures, which can help identify the most influential features in cancer detection.
- (3) *Adaptive artificial neural network (AANN)*: Adaptive artificial neural network, also known as adaptive neural network or adaptive neuro-fuzzy inference system (ANFIS), is a type of neural network that adjusts its structure and parameters during the learning process. AANN combines the capabilities of neural networks and fuzzy logic to create adaptive and interpretable models. AANN is particularly useful when dealing with complex and nonlinear relationships in cancer detection. It can adapt its network structure, such as the number of layers and neurons, and optimize its parameters to capture the underlying patterns in the data. AANN can provide interpretable rules based on fuzzy logic, which can aid in understanding the decision-making process.
- (4) *Back propagation neural network (BPNN)*: Back propagation neural network is a type of feed-forward neural network that uses the backpropagation algorithm to train the network. BPNN consists of multiple layers of interconnected neurons, including an input layer, one or more hidden layers, and an output layer. During

Table 4 Description of most commonly utilized features for machine learning algorithms

Types of features	Description
ABCD rule	The ABCD criterion evaluates the following four factors: A—asymmetrical, B—border, C—colours, and D—diameter
Autoregressive model (AM)	The autoregressive model (AM) is based on the presumption that the interaction between picture pixels takes place in a sequential fashion. According to this model, the illumination of the present pixel is dependent on the brightness of the pixels that came before it
Color	The overabundance of melanin beneath the outermost layer of the lesion is connected to color, and this excess causes a distinct color in the intensity of pixel values over a particular location. A pre-disposition for malignancy is suggested when there are numerous different hues present
Discrete cosines transform (DCT)	To facilitate the entry of data from a database, the discrete cosine transform (DCT) modifies the information in order that it has the potential to be embodied in basis functions, which correspond to the frequency components of the data
Discrete wavelet transform (DWT)	An effective method for extracting features from images is the discrete wavelet transform. Discrete wavelets transform (DWT) break down an image into smaller size components. In DWT, the energy of the image is represented as a function of frequency. Each resulting image is characterized by its energy, which is transformed again to larger sizes
First order features (FOF)	Without considering the spatial correlations of the pixel intensity distribution, first-order functions (FOF) generate image characteristics derived from image histogram analysis. The mean, variance, kurtosis, skewness coefficient, and other similar measures are all examples of FOF features
Gabor transform (GT)	The parameters obtained from Gabor Transform (GT) are frequency, variance, Gaussian envelope (σ), mean, orientation, magnitude, and standard deviation
Geometric	Some examples of Geometric features are solidity, area, size, equivalent diameter, convex area, perimeter, irregularity index radius, perimeter, area and circularity, bounding box dimensions, principal axis lengths, compactness, volume and elongation, Major axis, Minor axis length, Semi axis length, Solidity, Energy, Smooth, mean, Kurtosis, harmonic Mean
Gradient map features (GMF)	Evaluating the true magnitude of the luminance gradient at each location in a converted image, for instance, allows us to establish the statistical aspects of the spatial variation of the brightness distribution. This is what we call "gradient map features" (GMF)
Gray level co-occurrence matrix (GLCM)	The Gray Level Co-Occurrence Matrix (GLCM), which is an illustration of a radiomic and statistical feature extraction technique, shows the spatial interaction between pixels of various gray scales. Some of the most prominent GLCM-based features are Energy, Contrast, Homogeneity, Summation Average, Correlation, variance, Entropy, Differential Entropy, and the Second Moment
Gray-level size-zone matrix (GLSZM)	Quantifying gray level zones in a picture is another sort of radiomic feature that falls under this category. One such feature is known as the gray-level size-zone matrix (GLSZM). When there are a certain number of linked voxels that all have the same gray level intensity, we refer to this as a gray level zone
Grey-level run length matrix (GLRLM)	The homogenized run sizes are defined in the grey level run length matrix (GLRLM). Consecutive pixels with the same gray-level value make up a gray-level run. Gray level runs, which are the number of consecutive pixels with the same gray level value, are measured by this function
Haralick	Data pertaining to the "texture of images" is analyzed with the help of Haralick features. The Haralick texture feature allows for the computation of several parameters from a GLCM such as patch size, directionality, edge management, image quantization, and areas of interest
Histogram oriented of gradient (HOG)	HOG is an abbreviation for "histogram of oriented gradients," which refers to a histogram that is produced of the intensity of the gradient and direction around the image pixel. Quantification of the orientation is performed based on the number of angular bins, and within each bin, instances of angles with gradients are tallied. HOG features also offer an interpretation of the spatial pattern of variation in the luminance of the image
Hough transform (HT)	The form parameters may be estimated from the boundary points using a feature extraction approach in digital signal processing. This means that the Hough transform may be used to identify any kind of geometry
Hu invariant moments (HuIM)	The moment feature is used to broadly characterize by the geometric properties of the image region. The Hu invariant moments are a collection of seven integers that are independent of the specific transformation applied to the underlying image. Hu invariant moments are often utilized in the fields of image pattern recognition, target recognition, and classification, because they are invariant under rotation, translation, and scaling
Local Binary Patterns (LBP)	Local Binary Patterns (LBP) are a trustworthy texture-based feature that can extract the micro-texture properties of an image's content. The binary value is calculated by the thresholding technique of the eight adjacent pixels alongside the central pixel for every pixel in the image

Table 4 (continued)

Types of features	Description
Morphological	Morphology, in the field of medicine, is an exploration of the proportions and arrangement of microbes of various sizes and types. Anatomy may mean either gross or microscopic anatomy, although it most often describes the exploration of the latter. However, in common parlance, the two concepts are largely interchangeable
Radiomic	Radiomic features may reveal aspects of an illness that aren't immediately apparent to the naked human eye. The volume, shape, and surface, density and intensity, texture, tumor location, and relationships with surrounding tissues are only some of the many radiomic properties that may be assessed
Shape	Diameter, Area, Radius, Asymmetry index, Fractal dimension, Compactness index, Perimeter, and Circularity index are some examples of shape-based features
Statistical	The statistical features that are often used in order to derive mathematical details from a set of data include the mean, correlation, entropy, contrast, variance, inertia, standard deviation, energy, invariant moments, and homogeneity
Texture	The texture features extraction process involves collecting high-level data from an image, namely shape, color, contrast, and other features. Some of the examples of 2D texture features are, Mean, Variance, Skewness, Kurtosis, Energy, Entropy

training, the network adjusts the weights and biases of the neurons to minimize the difference between the predicted output and the true output. BPNN is known for its ability to learn complex mappings and capture intricate patterns in the data. It can handle both numerical and categorical data and has been widely used in various classification tasks, including cancer diagnosis. However, BPNN models are prone to overfitting and require careful tuning of hyperparameters.

- (5) *Radial basis function neural network (RBFNN)*: Radial basis function neural network is a type of neural network that uses radial basis functions as activation functions in its hidden layer(s). RBFNN is composed of three layers: an input layer, a hidden layer with radial basis functions, and an output layer. During training, the network adjusts the weights and biases of the neurons to minimize the difference between the predicted output and the true output. RBFNN is suitable for capturing complex and nonlinear relationships in cancer detection. The radial basis functions provide localized representations of the input space, allowing the network to model intricate patterns effectively. RBFNN is particularly useful when dealing with problems that require interpolation or approximation of unknown functions.
- (6) *Decision tree (DT)*: decision Tree is a simple yet powerful machine learning algorithm that uses a tree-like structure to make decisions based on feature values. Each internal node represents a decision based on a specific feature, and each leaf node represents a class label or a regression value. The decision tree is built by recursively splitting the data based on feature thresholds that optimize certain criteria, such as information

gain or Gini impurity. Decision trees are highly interpretable and can handle both numerical and categorical data. They can capture complex decision boundaries and provide insights into feature importance and interactions. However, decision trees are prone to overfitting, which can be mitigated by techniques like pruning or using ensemble methods.

- (7) *Ensemble classifier (EC)*: Ensemble classifier is a machine learning technique that combines the predictions of multiple base classifiers to make the final classification decision. The base classifiers can be any individual classifiers, such as decision trees, support vector machines, or neural networks. The combination of base classifiers can be done through voting, averaging, or weighted averaging. Ensemble classifiers aim to improve classification accuracy and generalization performance by reducing the bias and variance of individual classifiers. By leveraging the diversity of base classifiers, ensemble classifiers can capture different aspects of the data and make more robust predictions.

These classifiers, including SVM, RF, EC, AANN, BPNN, RBFNN, and DT, have been successfully applied in cancer detection tasks. Their effectiveness in accurately classifying and diagnosing various types of cancer can contribute to early diagnosis and timely treatment, ultimately improving patient outcomes. In summary, machine learning-based classifiers, including SVM, RF, EC, AANN, BPNN, RBFNN, and DT, have proven to be effective in cancer detection, particularly when combined with non-invasive diagnostic techniques.

Deep learning models

Deep learning is a branch of machine learning that focuses on training models to learn hierarchical representations of data directly from raw input. It has gained significant popularity in various fields, including cancer detection and segmentation, due to its ability to automatically learn complex patterns and features without the need for explicit feature engineering. In the context of cancer detection and segmentation, deep learning models have shown exceptional performance by leveraging large-scale biomedical images. Here is a detailed explanation of some commonly used deep learning models for these tasks:

- (1) *Convolutional neural networks (CNN)*: Convolutional neural networks (CNNs) are specifically designed for processing grid-like data, such as images. CNNs excel at capturing local patterns and spatial dependencies in an image. They employ convolutional layers that perform local receptive field convolutions, enabling the network to learn hierarchical features from the input image. These features capture different levels of abstraction, starting from low-level features like edges and textures to high-level features like shapes and objects. The learned features are then fed into fully connected layers for classification or further processing.

Popular CNN architectures used in cancer detection and segmentation tasks include:

- *AlexNet*: Introduced by Alex Krizhevsky et al., AlexNet was one of the pioneering deep CNN architectures. It consists of multiple convolutional and pooling layers, followed by fully connected layers.
- *ResNet*: Residual Network (ResNet) introduced the concept of residual connections, which allows for easier training of very deep neural networks. ResNet architecture includes skip connections that enable the gradient flow to bypass multiple layers, helping alleviate the vanishing gradient problem.
- *VGG*: The VGG network is known for its simplicity and uniform architecture. It consists of multiple convolutional layers with small receptive fields and pooling layers, leading to a deeper network.
- *UNet*: UNet is a widely used architecture for image segmentation tasks. It has an encoder-decoder structure with skip connections that preserve spatial information, making it effective for precise segmentation.

- (2) *Long short-term memory networks (LSTM)*: Long Short-Term Memory Networks (LSTMs) are a type of recurrent neural network (RNN) designed to model sequential and temporal data. LSTMs are particularly

useful when dealing with time-series data or sequences, such as analyzing longitudinal patient data or processing genomic sequences in cancer research. LSTMs have a memory cell that allows them to capture long-term dependencies and handle variable-length sequences effectively.

- (3) *Deep convolutional neural networks (DCNN)*: Deep Convolutional Neural Networks (DCNNs) are deep variants of CNNs with additional layers and parameters. DCNNs are capable of capturing complex spatial relationships in images. They often incorporate techniques like residual connections, dilated convolutions, or attention mechanisms to enhance their performance in cancer detection and segmentation tasks.
- (4) *Generalized adversarial networks (GAN)*: Generalized Adversarial Networks (GANs) consist of two components: a generator network and a discriminator network. GANs are trained in a competitive setting, where the generator network generates synthetic data samples, and the discriminator network aims to differentiate between real and synthetic data. GANs have been used for various tasks in cancer detection, such as generating realistic synthetic medical images or augmenting datasets to improve model performance.
- (5) *You only look once (YOLO)*: YOLO (You Only Look Once) is a deep learning model used for real-time object detection, including cancer detection in medical imaging. It can quickly and accurately identify and localize cancerous regions within medical images, such as mammograms or histopathology slides. YOLO works by dividing the input image into a grid and predicting bounding boxes and class probabilities for potential cancerous regions in each grid cell. It uses a single pass approach, eliminating the need for separate feature extraction steps. After predicting the bounding boxes, YOLO applies non-maximum suppression to remove redundant and overlapping boxes, leaving the most confident and non-overlapping detections. The output of the model includes the coordinates of the bounding boxes, class labels, and confidence scores for the detected cancerous regions. YOLO has different versions with improvements in accuracy, speed, and robustness. It has proven to be an effective tool for automating cancer detection and assisting medical professionals in analyzing medical images efficiently and accurately.
- (6) *Region-based convolutional neural networks (R-CNN)*: R-CNN is a deep learning model used for cancer detection. It involves generating potential regions of interest within an image, extracting features from these regions using a pre-trained convolutional neural network (CNN), and using support vector machines (SVMs) for classification and localization of cancerous regions.

R-CNN has been applied to various medical imaging modalities and can assist in early cancer diagnosis. It handles object detection and localization tasks but can be computationally expensive. Faster variants of R-CNN have been developed to improve efficiency. Overall, R-CNN is effective in automatically detecting and localizing cancerous regions within medical images, aiding in early detection and improved patient outcomes.

Researchers often customize existing deep learning architectures to suit specific cancer detection and segmentation tasks. These modifications may involve incorporating additional layers, adjusting network architecture, or introducing specialized techniques. For example, Optimized Convolutional Neural Networks (OCNN), Active Contour Convolutional Neural Networks (ACCNN), CNN + AlexNet, and Squeeze and Excitation-based CNN (SE-CNN) are customized CNN architectures used for cancer-related tasks. By leveraging deep learning models, these approaches can accurately and efficiently identify cancerous cells or tissues from medical images. This aids in early diagnosis, precise segmentation, and effective treatment planning, ultimately improving patient outcomes. These models contribute to advancing early cancer diagnosis, prognosis, and treatment strategies in the field of oncology.

Performance evaluation metrics (PEM)

In Machine learning and deep learning based classification task, the classifier performance measurement used to be done based on some performance evaluation metrics (PEM). Based on these metrics the any classification or detection task can be validated. There are various types of PEM that are being used by the researchers for validation of their proposed models but some of them are most common. Similarly in cancer detection task also there are various types of PEM are being utilized in various literatures, some of the mostly used metrics have been discussed in this study. The classification or detection of any task is mainly classified into two categories binary class (two-class) and multi-class (more than two). In binary classification, since there are only two classes, one class is called positive class and another one is called Negative class. Most of the PEMs can be calculated using confusion matrix or contingency table (Tharwat Aug. 2018). So, before understanding the PEM we must know about confusion matrix, the binary class confusion matrix is shown in Fig. 3. In confusion matrix, the actual outcome or prediction has been presented by green boxes and wrong outcomes or prediction in pink boxes. The key metrics include accuracy, error rate, specificity, sensitivity, precision, and F1-score, and their respective formulas are given in Eqs. (6–12). These metrics are commonly and widely used

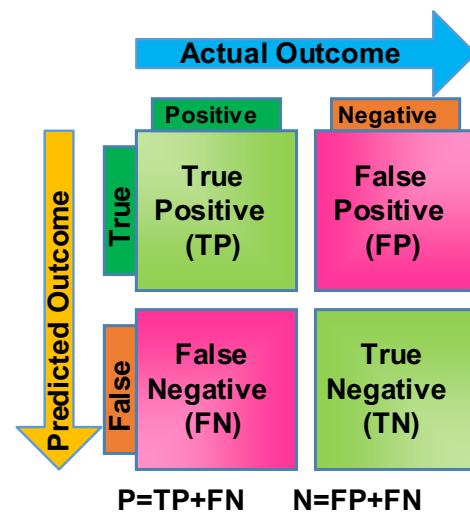


Fig. 3 Confusion matrix for binary class classification

in various classification problems and are formulated based on the confusion matrix.

$$\text{Accuracy (\%ACC)} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100 \quad (6)$$

$$\text{Errorrate (ER)} = 1 - \text{ACC} \quad (7)$$

$$\text{Errorrate (\%ER)} = \frac{FP + FN}{TP + TN + FP + FN} \times 100 = 100 - (\%ACC) \quad (8)$$

$$\text{Specificity (\%Sp)} = \frac{TN}{TN + FP} \times 100 \quad (9)$$

$$\text{Sensitivity (\%Sen)} = \frac{TP}{TP + FN} \times 100 \quad (10)$$

$$\text{Precision (\%Pr)} = \frac{TP}{TP + FP} \times 100 \quad (11)$$

$$\text{F1 - score (\%F1)} = \frac{2 \times TP}{(2 \times TP + FP + FN)} \times 100 \quad (12)$$

True positive (TP): Truly detected positive class, Positive class detected in positive categories. True negative (TN): Truly detected Negative class, Negative classes detected in negative categories. False positive (FP): Falsely detected positive class, negative class but falsely detected in positive categories. False negative (FN): Falsely detected negative class, positive class but falsely detected in negative categories. Total positive = True positive (TP) + False negative (FN). Total negative = True negative (TN) + False positive (FP). True classification = True positive (TP) + False positive (FP).

(FP). False classification = False negative (FN) + True negative (TN).

Cancer detection analysis

In this section, we will discuss the diagnosis and detection methods for different types of cancer using machine learning (ML) and deep learning (DL) algorithms. Specifically, we will focus on brain, lung, skin, and breast cancer and provide a visual representation of these cancer types. For each type of cancer, we will analyze and review literature that utilizes traditional machine learning techniques. This analysis includes information such as the authors and publication year, the features used in cancer detection, the best machine learning models employed, the dataset used for cancer detection, and the highest accuracy achieved by the best model. We will also present a state-of-the-art table that separately compares the use of traditional machine learning algorithms and deep learning algorithms for cancer detection. Both ML and DL models have distinct structural and functional differences, so we will analyze and discuss them separately to provide a comprehensive understanding of their capabilities and performance in cancer detection. By examining the literature and conducting a comparative analysis, we aim to highlight the advancements and achievements in cancer detection using ML and DL algorithms. This information can contribute to the development of more accurate and effective cancer diagnostic tools and improve patient outcomes.

Brain tumor detection

Brain tumor

The brain is the most sophisticated and important part of the human body which controls the entire functioning of actions and activities. Controlling blood circulation to breathing, even movement of any body part, and most importantly thoughts

are being controlled through the human brain. So, it is very important for all of us that the brain should function properly and any malignancy in the brain should be diagnosed as early as possible. As we know that tumor which are primarily classified into two classes benign and malignant, where the benign tumor does not have tendency to spread over the other body part whereas the malignant tumors are very dangerous, and they have tendency to spread over the neighboring cell even to other human body parts. A benign tumor is not cancerous in nature, and it is not very alarming if it is in the other body parts except brain. So, even a benign tumor in the human brain can be as dangerous as the malignant tumor in other body parts hence it must be detected at a very early stage to save the life of the person suffering from it. As per the report published in Indian express, around 40–50 thousand peoples diagnosed with brain cancer among which 20% are children of age less than 18 years (The Indian Express 2019). The survival rate in case of brain tumor is much less as compared to other diseases, which makes it very lethal and deadly than other cancer types. Some other modalities are also available for human brain cancer diagnoses along with MRI scans are PET, CT, fMRI, and mpMRI. The brain scanning technique comes under a non-invasive method which is painless and affordable and available without surgery. The MRI scans of human brain can be further categorized as FLAIR, T1 weighted, T2 weighted, and T1 contrast images. The pictorial visualization of four different types of MRI scans, FLAIR, T1 weighted, T2 weighted, and T1 contrast, respectively, is presented in Fig. 4.

Brain tumor detection using ML

In this section, we have reviewed and presented a state-of-art table for the detection of brain tumor using traditional machine learning methods. We have included the recently published research articles during 2018–2023 in the review table for the ML-based cancer detection techniques. We have only considered the highest accuracy (%) as the performance evaluation metrics to maintain the homogeneity for

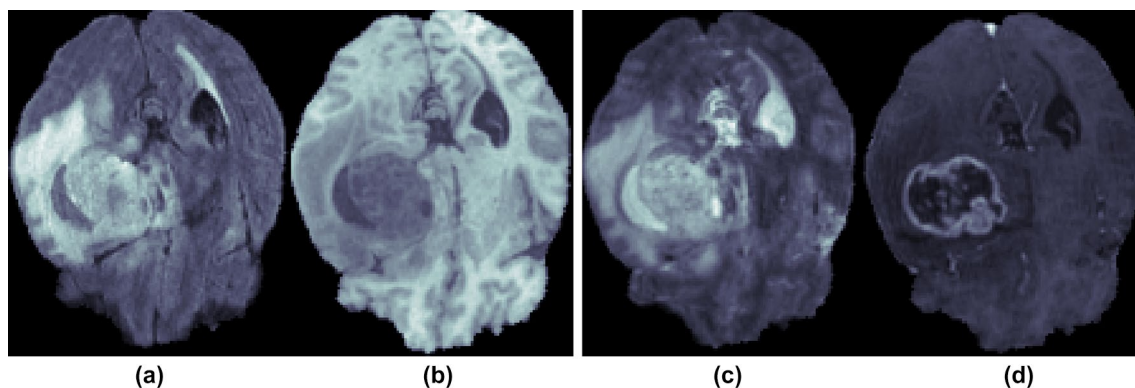


Fig. 4 Pictorial view of **a** FLAIR, **b** T1 Weighted, **c** T2 weighted, **d** T1 Contrast MRI scans of brain

Table 5 Machine Learning based brain tumor detection techniques

Author/year	Literature	Features	Proposed model	Dataset/images	Highest accuracy (%)
Al-Dmour and Al-Ani 2018	Al-Dmour and Al-Ani (2018)	SA, Pixel Intensity	CFNN	IBSR18, IBSR20	98.10
Gupta et al. 2019	Gupta et al. (2019)	GLCM, LBP, GLRLM, PHOG	EC	BRATS, JMCD	97.76
Salem Ghahfarrokhi and Khodadadi 2020	Salem Ghahfarrokhi and Khodadadi (2020)	GLCM, DWT	PN	3064 MRI	98.90
Virupakshappa and Amara-pur 2020	Virupakshappa and Amara-pur (2020)	MWD, GLCM, GT, Moment Invariant	AANN + WOA	BRATS 2015	98.00
Brunese et al. 2020	Brunese et al. (2020)	Radiomic Features (FOF, Shape, GLCM, GLRLM, GLSZM)	EC	REMBRANDT, BraTS 2019, Radiopaedia (111,205 MRI)	99.00
Shafi et al. 2021	Shafi et al. (2021)	GLCM, GLRLM, GLSZM, NGTDM	EC	2399 MRI	97.96
Vankdothu and Hameed 2022	Vankdothu and Hameed (2022)	GLRLM	SVM	210 CT	99.24
Malarvizhi et al. 2022	Malarvizhi et al. (2022)	DWT, GLCM	SVM	15 MRI	80.00
Shimanto et al. 2023	Shimanto et al. (2023)	Texture, GLCM	ADSVM	70 MRI	97.14
Archana and Komarasamy 2023	Archana and Komarasamy (2023)	Segmented features	BKNN	3,064 MRI	97.70
Rasheed et al. 2023	Rasheed et al. (2023)	GLCM-based Statistical feature	SVM	3541 MRI	98.00

Pyramid of Histogram of Gradients (PHOG), Pattern Net (PN), Adaptive Artificial Neural Network (AANN), Whale Optimization Algorithm (WOA), Superpixel Algorithm (SA), Anisotropic Diffusion based SVM (ADSVM), Bagging Ensemble with K-Nearest Neighbor (BKNN), Diagnosis Centre at NSCB Medical College Jabalpur (JMCD), Internet Brain Segmentation Repository (IBSR), Magnetic Resonance Imaging (MRI)

the validation of utilized classifier models. Table 5 presents the state-of-art table for the brain cancer detection technique using traditional machine learning models only.

The best performances and worst performances in terms of accuracy for brain tumor detection techniques, along with some details about each study has been presented based on Table 5.

Best performances

- Vankdothu and Hameed (2022): This study achieved an accuracy of 99.24% by employing the Gray Level Run Length Matrix (GLRLM) feature and Support Vector Machine (SVM) as the classification model. The dataset used in this study included 210 CT scans for brain tumor detection.
- Brunese et al. (2020): This study achieved a high accuracy of 99.00% using radiomic features such as First Order Features (FOF), Shape, Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), and Gray Level Size Zone Matrix (GLSZM). The proposed model was evaluated on multiple datasets including REMBRANDT, BraTS 2019, and Radiopaedia, which consisted of a total of 111,205 MRI images.

- Salem Ghahfarrokhi and Khodadadi (2020): With a high accuracy of 98.90%, this study utilized features such as Gray Level Co-occurrence Matrix (GLCM) and Discrete Wavelet Transform (DWT). The proposed model was trained and evaluated on a dataset consisting of 3,064 MRI images.

Worst performances

- Malarvizhi et al. (2022): This study achieved an accuracy of 80.00% using features like Discrete Wavelet Transform (DWT) and Gray Level Co-occurrence Matrix (GLCM) with Support Vector Machine (SVM) as the classification model. The dataset used for evaluation included only 15 MRI images, which might have limited the generalizability of the results.

Key findings

- Machine learning techniques have shown high accuracy in brain tumor detection, ranging from 80 to 99.24%.
- Various features, such as statistical features, GLCM, GLRLM, GLSZM, and texture features, contribute to accurate detection.

- c. Different models, including CFNN, SVM, AANN + WOA, and BKNN, have been proposed for brain tumor detection.
- d. Datasets such as BRATS, JMCD, IBSR18, IBSR20, REMBRANDT, and BraTS 2019 have been used for training and evaluation.

Challenges

- a. Limited availability of labeled datasets with diverse brain tumor types and imaging modalities.
- b. Robustness of models in handling variations in image quality, noise, and artifacts.
- c. Interpretability of machine learning models to aid in clinical decision-making.
- d. Integration of models into existing healthcare systems for real-time tumor detection.
- e. Addressing ethical concerns related to data privacy and security.

Brain tumor detection using DL

As deep learning model-based detection techniques are more commonly and majorly used by most of the researchers from past few years, we have separately considered reviewing DL-based cancer detection techniques. We have surveyed the peer reviewed work published between year 2018 and 2023 and those have utilized the DL models for brain cancer detection. In this review, we have excluded works that do not utilize classification accuracy as the

evaluation metric. Additionally, literature that employed both ML and DL classifiers for brain tumor detection was excluded, focusing solely on studies that utilized DL models with accuracy as the detection metric. Table 6 provides a comprehensive overview of deep learning-based brain tumor detection techniques, exclusively considering accuracy as the evaluation criterion.

We have presented the analysis of best performances and worst performances for brain tumor detection based on deep learning models listed in Table 6.

Best performances

- a. Krishnapriya and Karuna (2023): CNN (VGG-19) achieved the highest accuracy of 99.48%. The model was trained on a dataset of 305 MRI images. The VGG-19 architecture was employed, which is a deep convolutional neural network known for its excellent performance in image classification tasks. This technique demonstrated superior accuracy in brain tumor detection, making it one of the top-performing approaches.
- b. Ramtekkar et al. (2023): GLCM-based OCNN achieved an accuracy of 98.90% using a dataset of 253 MRI images. GLCM (Gray Level Co-occurrence Matrix) features were extracted to capture textural properties of the brain tumor regions. The OCNN (One-Class Neural Network) technique, known for its anomaly detection capabilities, was utilized to classify tumor and non-tumor regions. This approach demonstrated high accuracy in identifying brain tumors.

Table 6 Deep Learning based brain tumor detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
Shahzadi et al. 2018	Shahzadi et al. (2018)	VGG-16	CNN-LSTM	60 MRI	84.00
Talo et al. 2019	Talo et al. (2019)		CNN	1074 MRI	95.23
Toğaçar et al. 2020	Toğaçar et al. (2020)		BrainMRNet	253 MRI	96.05
Safdar et al. 2020	Safdar et al. (2020)		YOLO v3	1961 MRI	96.00
Ghassemi et al. 2020	Ghassemi et al. (2020)		GAN + CNN	3064 MRI	95.60
Hashemzahi et al. 2020	Hashemzahi et al. (2020)		CNN-NADE	3064 MRI	95.00
Çinar and Yildirim 2020	Çinar and Yildirim (2020)		Hybrid CNN	253 MRI	97.20
Karayegen and Aksahin 2021	Karayegen and Aksahin (2020)		CNN	257 MRI	95.70
Shetty et al. 2022	Shetty et al. (2022)		CNN	1500 MRI	98.21
Ramtekkar et al. 2023	Ramtekkar et al. (2023)		GLCM	253 MRI	98.90
Saeedi et al. 2023	Saeedi et al. (2023)	Shape, Intensity, Model	2D CNN	3264 MRI	96.47
Mahmud et al. (2023)	Mahmud et al. (2023)	SCNN	CNN	3264 MRI	93.30
Patil et al. (2023)	Patil and Kirange (2023)		EDCNN	3064 MRI	97.77
Krishnapriya and Karuna 2023	Krishnapriya and Karuna (2023)		CNN (VGG-19)	305 MRI	99.48

Neural Autoregressive Distribution Estimation (NADE), Ensemble Deep Convolutional Neural Network (EDCNN), Shallow Convolutional Neural Network (SCNN)

Worst performances

- a. Shahzadi et al. (2018): In the study, combination of VGG-16 and CNN-LSTM techniques was utilized for breast cancer detection using MRI images. The VGG-16 model was employed to extract features from the images, which were then fed into a CNN-LSTM architecture to capture spatial and temporal dependencies. The proposed approach achieved a best accuracy of 84.00% in classifying breast cancer cases in a dataset of 60 MRI images. Further research is needed to validate the effectiveness of this method on larger datasets.

Key findings

- a. VGG-16, CNN, and BrainMRNet models showed promising results with accuracies of 84.00%, 95.23%, and 96.05% respectively.
- b. Techniques such as YOLO v3, GAN + CNN, and CNN-NADE achieved accuracies of 96.00%, 95.60%, and 95.00% respectively.
- c. Hybrid CNN and LBP-CNN models demonstrated good performance with accuracies of 97.20% and 98.21% respectively.
- d. GLCM-based models, including OCNN and GLCM-CNN, showed high accuracies of 98.90%.
- e. Ensemble Deep CNN (EDCNN) achieved an accuracy of 97.77% using the SCNN technique.
- f. CNN models, including VGG-19, achieved excellent accuracy of 99.48%.

Challenges

- a. *Limited dataset*: Some studies utilized small datasets, which might affect the generalization of the models. Larger and diverse datasets are required for better evaluation and robust performance.
- b. *Model complexity*: Deep learning models can be complex, requiring significant computational resources and time for training. Optimizing and fine-tuning these models can be challenging.
- c. *Interpretability*: Deep learning models often lack interpretability, making it difficult to understand the reasoning behind their predictions. Interpretable models are essential in the medical field to gain trust and provide actionable insights.
- d. *Class imbalance*: Imbalanced class distribution, where certain tumor types are underrepresented, can impact the model's performance and bias the results. Proper data handling techniques are necessary to address class imbalance.
- e. *Dataset variability*: Variations in MRI acquisition protocols, image resolutions, and tumor characteristics across different datasets can introduce challenges in developing

robust models that generalize well across diverse data sources.

- f. *Deployment and validation*: The translation of deep learning models from research to clinical practice requires rigorous validation, addressing regulatory requirements, and ensuring reliable and safe deployment.

Lung cancer detection

Lung cancer

Lung cancer is a type of cancer that starts in the lungs and can spread to other parts of the body. It is one of the most common types of cancer, and it is the leading cause of cancer-related deaths worldwide. Lung cancer can be broadly classified into two main types: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) (Mokoatle et al. 2023). NSCLC is the most common type of lung cancer, accounting for about 85% of all cases. It includes three subtypes: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. SCLC is a less common type of lung cancer, accounting for about 15% of all cases. It is more aggressive and tends to spread more quickly than NSCLC. According to the World Health Organization (WHO), lung cancer is the most commonly diagnosed cancer worldwide, with an estimated 2.1 million new cases in 2018. It is also the leading cause of cancer-related deaths, accounting for 1.8 million deaths in 2018. The incidence and mortality rates of lung cancer vary by region, with the highest rates observed in developed countries. Early detection of lung cancer is crucial for improving patient outcomes. Studies have shown that early detection through screening can reduce mortality rates from lung cancer. This is because early-stage lung cancer is more treatable than advanced-stage lung cancer. Screening for lung cancer is recommended for individuals who are at high risk, such as current or former smokers. The most common screening modality is low-dose CT scan, which can detect lung cancer at an early stage when it is more treatable. Figure 5 displays some sample CT images for the diagnosis of lung cancer that were taken from the LIDC/IDRI Lung Cancer Dataset.

Lung cancer detection using ML

Lung cancer detection or classification using traditional machine learning methods has been presented in the state-of-art Table 7. The table includes information about the author and year of the study, the features used in the model, the technique used to develop the model, the dataset/images used for training and testing, and the best accuracy achieved by the model. The table includes 15 studies related to the use of various techniques and features for detecting lung cancer. The studies were conducted between 2018 and 2023. The

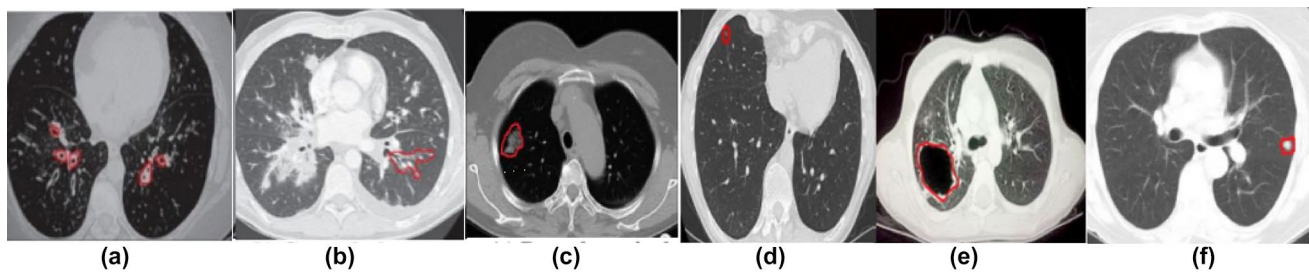


Fig. 5 An example of images taken from LIDC/IDRI dataset of classification of six lung disease patterns, **a** honeycombing, **b** ground-glass, **c** bronchovascular, **d** nodular, **e** emphysema-like, **f** normal

Table 7 Machine Learning based lung cancer detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
Woźniak et al. 2018	Woźniak et al. (2018)	Momentum, Contrast	PNN	100 X-ray	92.00
Chang et al. 2018	Chang et al. (2018)	PCA, Sensor Data	MLP	85 volunteers	75.00
Zhou et al. 2018	Zhou et al. (2018)	Radiomic features, GLCM, GLRLM, GT, WT	SVM	348 CT	85.98
Wang et al. 2018	Wang et al. (2018)	Haralick, GLCM, Morphological	ELM	1018 CT	95.91
Radhika et al. 2019	Radhika et al. (2019)	23 Lung Cancer Attributes	SVM	UCI	99.20
Khan et al. 2019	Khan et al. (2019)	GLCM, LBP, Color, DCT	SVM	DermIS	96.00
Naqi et al. 2019)	Naqi et al. (2019)	HoG, PCA, Texture, Geometry	AdaBoost	LIDC	99.20
Akkar and Haddad 2020	Akkar and Haddad (2020)	Histogram, GLCM	BPNN	100 Images	95.918
Khan et al. 2020	Khan et al. (2020)	DWT, Geometric, MSER, HGLCM, LBP	EC	LDSBD-2017	99.40
Bębas et al. 2021	Bębas (2021)	Texture, FOF, GMF, GLRLM GLCM, AM, GT, LBP, HOG, DWT	SVM	155 MRI	75.48
Nanglia et al. 2021	Nanglia et al. (2021)	SURF	KASC	ELCAP-PLID	98.08
Shim et al. 2022	Shim et al. (2022)	Radiomic, Statistical, Haralick texture, GLCM, GLSZM, GLRLM	SVM	LCA	99.89
Mezher et al. 2022	Mezher et al. (2022)	Lung Cancer Attributes	SVM	309 Images	96.20
Prakash et al. 2023	Prakash et al. (2023)	GLCMWAA	EESNN	IQ-OTHNCCD/1190 Images	99.12
Nigudgi and Bhyri 2023	Nigudgi and Bhyri (2023)	AlexNet, VGG and GoogleNet	SVM	IQ-OTHNCCD/1190 Images	97.00

ELCAP Public Lung Image Database (ELCAP-PLID), Maximally Stable External Regions (MSER), Harlick GLCM (HGLCM), Speeded Up Robust Features (SURF), GLCM Window Adaptive Approach (GLCMWAA), Probabilistic Neural Network (PNN)

techniques used in the studies include PNN, MLP, SVM, ELM, AdaBoost, BPNN, EC, and EESNN. The best accuracy percentage achieved in the studies ranges from 75.48 to 99.89%. The most commonly used technique was SVM, which was used in seven out of the 15 studies. The GLCM feature is the most used feature for lung cancer detection, followed by radiomic features and Haralick.

The detailed explanations of the best and worst performances in terms of accuracy for the detection of lung cancer using ML models are presented in Table 7.

Best performances

- Shim et al. (2022) achieved the highest accuracy of 99.89% using SVM with radiomic, statistical, Haralick texture, GLCM, GLSZM, and GLRLM features. The study used the Lung Cancer Atlas (LCA) dataset consisting of 1,363 CT images. The high accuracy achieved in this study can be attributed to the use of a combination of different feature types and the large dataset used.
- Khan et al. (2020) achieved an accuracy of 99.40% using EC with DWT, geometric, MSER, HGLCM, and LBP

features. The study used the Lung Dataset from Shaikat Khanum Memorial Cancer Hospital and Research Centre (SKMCH&RC) consisting of 1,190 images. The use of multiple feature types and the relatively large dataset contributed to the high accuracy achieved in this study.

- c. Radhika et al. (2019) achieved an accuracy of 99.20% using SVM with 23 lung cancer attributes. The study used the UCI dataset consisting of 32 instances. The high accuracy achieved in this study can be attributed to the use of a large number of attributes and the relatively small dataset used.

Worst performances

- a. Bębas (2021) achieved an accuracy of 75.48% using SVM with texture, FOF, GMF, GLRLM, GLCM, AM, GT, LBP, and HOG features. The study used the MRI dataset consisting of 155 images. The low accuracy achieved in this study can be attributed to the relatively small dataset used and the use of a large number of feature types.
- b. Chang et al. (2018) achieved an accuracy of 75.00% using MLP with PCA and sensor data. The study used a dataset consisting of sensor data from 85 volunteers. The low accuracy achieved in this study can be attributed to the small dataset used and the use of only one feature type.

Overall, the observation shows that the range of accuracies achieved in the table is quite wide, with the best accuracy being almost 25 percentage points higher than the worst accuracy. This suggests that the effectiveness of the models developed in these studies varies widely, and that some models are much more effective than others at detecting lung cancer. The features used in each study, as well as the size and quality of the datasets used, likely play a significant role in determining the accuracy of the models. Therefore, it is important to carefully choose the appropriate features and techniques for the specific dataset and problem at hand.

Key findings

- a. The table summarizes various studies that have used machine learning techniques for lung cancer detection.
- b. The studies have used different feature types, techniques, and datasets/images.
- c. The best accuracy achieved ranges from 75.48 to 99.89%.
- d. The top 3 best performances are achieved using SVM with a combination of different feature types, while the worst 3 performances are achieved using a relatively small number of feature types or a small dataset.

- e. The studies show that machine learning techniques have the potential to improve the accuracy of lung cancer detection.

Challenges

- a. One of the main challenges in using machine learning techniques for lung cancer detection is the availability of high-quality datasets.
- b. Many studies have used small datasets, which can limit the generalizability of the results.
- c. Another challenge is the robustness of the models.
- d. Machine learning models can be sensitive to changes in the input data, which can affect their performance.
- e. Additionally, the interpretability of the models can be a challenge.
- f. Machine learning models can be complex, and it can be difficult to understand how they make predictions.
- g. Finally, there is a need to integrate machine learning techniques into clinical practice.
- h. This requires addressing ethical, regulatory, and technical challenges.

Lung cancer detection using DL

Table 8 shows a list of deep learning-based techniques for lung cancer detection, along with their features, techniques, dataset/images used, and the best accuracy achieved. The studies in the table used various deep learning techniques, such as CNN, DenseNet, UNet, ResNet, and CapsNet, among others. The datasets used in these studies include LIDC-IDRI, TCIA, LUNA16, and Lung Cancer Dataset, among others. Some studies used a combination of multiple datasets to improve the accuracy of their models. The features used in these studies include taxic weights, morphological patterns, statistical and multi-scale features, and deep features, among others. Some studies used a combination of multiple features to improve the accuracy of their models. The accuracy achieved by these studies ranges from 74.40 to 99.76%. Some studies achieved high accuracy rates of over 98%, demonstrating the potential of deep learning-based techniques for lung cancer detection.

Below is the analysis of the best and worst performances for the detection of lung cancer using DL models, based on accuracy as listed in Table 8.

Best performances

- a. Alsheikhy et al. (2023): In their study, Alsheikhy et al. achieved the highest accuracy of 99.76% for lung cancer detection. They used deep features extracted from VGG-19 combined with LSTMs (long short-term memory) to classify Chest CT Scans, LIDC-IDRI, and LUNA16

Table 8 Deep Learning based lung cancer detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
de Carvalho Filho et al. 2018	Carvalho Filho et al. (2018)	Taxic Weights, PhT	CNN	LIDC-IDRI	92.60
Ausawalaithong et al. 2019	Ausawalaithong et al. (2018)	Histogram Equalization	DenseNet	JSRT, CX-14	74.40
Saric et al. 2019	Saric et al. (2019)	ROI	CNN	ACDC@LUNGHP	97.90
Bhatia et al. 2019	Bhatia et al. (2019)	UNet, ResNet	XGBoost + RF	LIDC-IDRI	84.00
Asuntha and Srinivasan 2020	Asuntha and Srinivasan (2020)	HoG, LBP, SIFT, Zernike Moment	FPSO-CNN	LIDC-IDRI	95.60
Yu et al. 2020	Yu et al. (2020)	Morphological Patterns	CNN	TCGA, ICGC	90.00
Masud et al. 2021	Masud et al. (2021)	DFT, DWT	CNN	LC25000 Images	96.30
Faruqui et al. 2021	Faruqui et al. (2021)	Shape	LungNet	LIDC-IDRC	96.81
Selvapandian et al. 2022	Selvapandian et al. (2022)	Statistical, Multi-scale	SCSF-based GAN	LIDC, TCIA	98.51
Vaiyapuri et al. 2022	Vaiyapuri et al. (2022)	NASNetLarge	CSO-CADLCC	300 CT images	98.89
Feng and Jiang 2022	Feng and Jiang (2022)		Mask-RCNN and DPN	90 CT/MRI	97.94
Naseer et al. 2022	Naseer et al. (2022)	CNN	CNN-AlexNet	LUNA16	97.25
Kasinathan and Jayakumar 2022	Kasinathan and Jayakumar (2022)		M-CNN	LIDC-IDRI/PET	97.10
Ruan et al. 2022	Ruan et al. (2022)		DL	1984 CT scans	97.08
Pradhan et al. (2023)	Pradhan et al. (2023)	PCA, T-SNE, CNN	BF-SSA-HR-DEL	Lung Cancer Dataset	93.15
Kumar et al. (2023)	Kumar et al. (2023)	U-Net	CNN	Lung CT Images	97.10
Alsheikhy et al. (2023)	Alsheikhy et al. (2023)	Deep Features	VGG-19 + LSTMs	Chest CT Scans, LIDC-IDRI, LUNA16	99.76
Heidari et al. (2023)	Heidari et al. (2023)		CapsNet + Federated	CIA, KDSB, LUNA 16, Local	99.69

Japanese Society of Radiological Technology(JSRT), Chest X-ray14 (CX-14), International Cancer Genome Consortium (ICGC), sine cosine Sail Fish (SCSF), Phylogenetic Trees (PhT), Scale Invariant Feature Transform (SIFT), Discrete Fourier Transform (DFT), cat swarm optimization-based computer-aided diagnosis model for lung cancer classification (CSO-CADLCC), Dual Path Network (DPN), Positron Emission Tomography (PET), Fuzzy Particle Swarm Optimization (FPSO), Region of Interest (ROI), multilayer convolutional neural network (M-CNN), t-Distributed Stochastic Neighbor Embedding (T-SNE), Best Fitness-based Squirrel Search Algorithm (BF-SSA), High Ranking Deep Ensemble Learning (HR-DEL)

datasets. This approach demonstrated exceptional accuracy in identifying lung cancer cases.

- b. Heidari et al. (2023): The 2nd highest accuracy of 99.69% was achieved in their study. They utilized the CapsNet (capsule network) combined with Federated Learning approach for lung cancer detection. The study employed multiple datasets including CIA, KDSB, LUNA 16, and Local, and achieved an impressive accuracy of 99.69%.
- c. Vaiyapuri et al. (2022): Vaiyapuri et al. achieved a best accuracy of 98.89% for lung cancer detection. They utilized the NASNetLarge architecture with a CSO-CADLCC (Cross-Sequence Optimization for Convolutional Neural Networks with Cascaded AdaBoost, Discriminative Learning, and Closed-Loop Classification) technique. The study used a dataset consisting of 300 CT images to train and validate their model, resulting in high accuracy.

Worst performances

- a. Ausawalaithong et al. (2018): Ausawalaithong et al. achieved the lowest accuracy of 74.40% for lung cancer detection. Their approach involved using histogram equalization in combination with DenseNet, a deep convolutional neural network. The study used the JSRT (Japanese Society of Radiological Technology) and CX-14 datasets, but the accuracy achieved was relatively lower compared to other techniques.
- b. Bhatia et al. (2019): Bhatia et al. reported an accuracy of 84.00% for lung cancer detection. They employed a combination of UNet and ResNet models along with XGBoost (Extreme Gradient Boosting) and Random Forest classifiers. The study utilized the LIDC-IDRI dataset for training and evaluation, but the accuracy achieved was relatively moderate compared to other methods.

Key findings

- Deep learning techniques have been successfully applied to detect lung cancer.
- The best accuracy achieved is 99.76% using VGG-19 + LSTMs with chest CT scans, LIDC-IDRI, and LUNA16 datasets.
- Most of the techniques use CNN as the primary model for lung cancer detection.
- The datasets used for training and testing the models include LIDC-IDRI, LUNA16, TCGA, ICGC, LC25000 Images, and Lung CT Images.
- The features used for detection include Taxic Weights, PhT, Histogram Equalization, Morphological Patterns, DFT, DWT, and Statistical and Multi-scale features.

Challenges

- The availability of high-quality annotated datasets for training the models.
- The requirement of large amounts of computational resources for training and testing the models.
- The need for specialized expertise in deep learning techniques to develop accurate models.

- The challenge of integrating the developed models into the clinical workflow for real-time diagnosis and treatment planning.

Skin cancer detection

Skin cancer

Skin cancer is a prevalent disease that can develop on human skin due to abnormal growth of skin tissues caused by exposure to ultraviolet radiation from the sun. The World Health Organization predicts that the number of skin cancer cases will rise to 2.2 million by 2025. Basal cell carcinoma, squamous cell carcinoma, melanoma, and actinic keratosis are some of the most common forms of skin cancer, with melanoma being the most deadly and responsible for 75% of skin cancer deaths. Early diagnosis of skin cancer is crucial, and studies have shown that diagnosing skin cancer at an earlier stage can increase a person's chance of surviving the disease by 91%. Machine learning (ML) and deep learning (DL) techniques can be used to identify various types of skin cancer from a variety of sources. These techniques can analyze images of skin lesions and identify patterns that may indicate the presence of cancer. By training ML and DL models on large datasets of skin lesion images, these models can become highly accurate at identifying different types

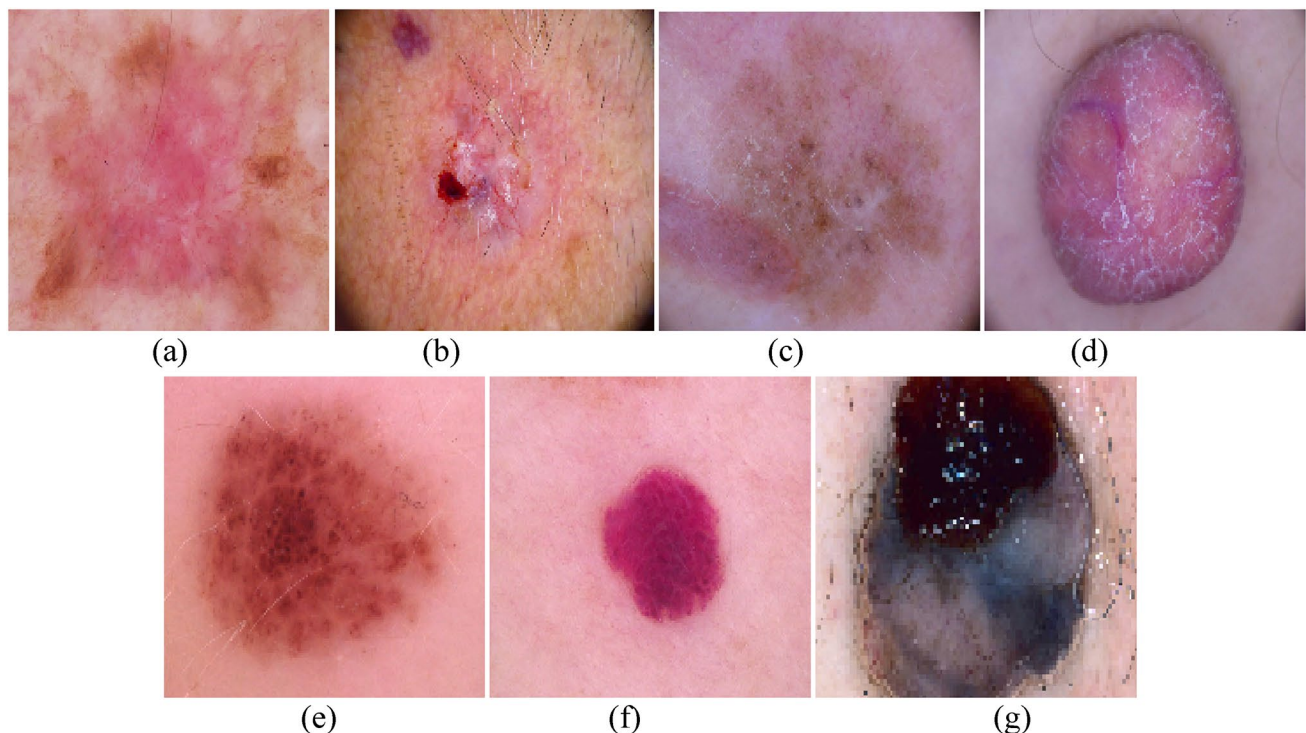


Fig. 6 Sample skin cancer images from the HAM10000 dataset. The images show **a** actinic keratosis, **b** basal cell carcinoma, **c** benign keratosis-like lesions, **d** dermatofibroma, **e** melanocytic nevi, **f** vascular lesions, and **g** melanoma

of skin cancer. In this paper, we discuss how ML and DL techniques can be used to identify skin cancer from a variety of sources, including the HAM10000 dataset, to improve early detection and ultimately save lives. Figure 6 shows sample skin cancer images from the HAM10000 dataset. The images depict different types of skin lesions, including actinic keratosis, basal cell carcinoma, benign keratosis-like lesions, dermatofibroma, melanocytic nevi, vascular lesions, and melanoma. The caption lists each type of skin lesion and assigns a letter to each corresponding image, ranging from (a–g).

Skin cancer detection using ML

Table 9 provides a summary of various machine learning-based skin cancer detection techniques published in recent years. The table includes the author/year of publication, literature reference, features used, machine learning technique employed, dataset/images used, and the best accuracy achieved by each technique. The table includes 17 studies that used different features, techniques, and datasets to train their ML models. The techniques listed in the table use

different sets of features, including the ABCD rule, shape, texture, color, and statistical features. The machine learning techniques employed include artificial neural networks, support vector machines, k-nearest neighbours, and random forests, among others. The datasets used in these studies include PH², ISIC, HAM10000, and others. The best accuracy achieved by these techniques ranges from 80 to 99.9%, with some achieving accuracy rates close to or above 98%. These results demonstrate the potential of machine learning models in accurately detecting skin cancer and the importance of using large and diverse datasets to train these models.

The analysis of some best-performing and worst-performing skin cancer detection techniques from Table 9 is presented below.

Best performances

- a. Wahba et al. (2018): This study achieved the highest accuracy of 99.90% using GLDM-CLDM (Gray-Level Difference Method-Completed Local Descriptor Matrix) and the ABCD (Asymmetry, Border irregularity, Color variegation, and Diameter) rule as features with the

Table 9 Machine learning based skin cancer detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
Wahba et al. 2018	Wahba et al. (2018)	GLDM-CLDM, ABCD	Cubic SVM	ISIC-2018/1200 Images	99.90
Murugan et al. 2019	Murugan et al. (2019)	Shape, ABCD Rule, GLCM	KNN, RF, SVM	ISIC /1000 Images	89.43
Chatterjee et al. 2019	Chatterjee et al. (2019)	Morphological, BIM, Texture (GLCM, FRTA) Color	SVM-RFE	ISIC, PH ² , ATLAS, IDS	95.21
Iqtidar et al. 2020	Iqtidar et al. (2020)	LBP, HOG, BoVW	KMC	HAM 10000	99.80
Monika et al. 2020	Monika et al. (2020)	ABCD Rule, GLCM	MSVM	ISIC 2019/800 Images	96.25
Amin et al. 2020	Amin et al. (2020)	AlexNet, VGG16	EC	PH ² , ISBI/7849 Images	99.90
Senan and Jadhav 2021	Senan and Jadhav (2021)	ABCD	TDS	PH ²	84.00
Tumpa and Kabir 2021	Tumpa and Kabir (2021)	ABCD, GLCM, LBP	FFNN-BP	PH ²	97.70
Bi et al. 2021)	Bi et al. (2021)	Geometric, Texture, Statistical	SVM	ACS, ISIC	87.50
Bhargavi et al. 2022	Bhargavi et al. (2022)	Hough Transform	SVM	ISIC	80.00
Ahammed et al. 2022	Ahammed et al. (2022)	GLCM, Statistical Features	SVM	ISIC, HAM10000	97.00
Batista et al. 2022	Batista et al. (2022)	Texture, Color, Generic	BIC + RF	HAM10000/10,015 Images	93.00
Zeng et al. 2023	Zeng et al. (2023)	Min.Value, SE, PE, ShE	PSO-SVM	150 Samples	97.78
Jaculin Femil and Jaya 2023	Jaculin Femil and Jaya (2023)	GRCM	PNN	PH ² , Dermis, Dermquest, Kaggle	97.83
Keerthana et al. 2023	Keerthana et al. (2022)	Pre-trained CNN	SVM	ISIC, ISBI 2016	88.02
Maqsood and Damaševičius 2023	Maqsood and Damaševičius (2023)	Deep feature	MC-SVM	HAM10000, ISIC18, ISIC19, PH ²	98.98

Cumulative Level-Difference Mean (CLDM), Gray-Level Difference Method (GLDM), recursive feature elimination (RFE), Fractal Texture Analysis (FRTA), Border irregularity measure (BIM), Automatic Tool for Local Annotations and Segmentation (ATLAS), International Dermoscopy Society (IDS), Feedforward Neural Network with Backpropagation (FFNN-BP), Border-Interior Classification (BIC), Gabor Response Co-occurrence Matrix (GRCM), multi-class SVM (MC-SVM)

Cubic SVM technique. The ISIC-2018 dataset, consisting of 1,200 images, was utilized. The combination of texture-based features and a powerful classification technique contributed to the outstanding accuracy achieved.

- b. Amin et al. (2020): Similar to the previous analysis, this study also achieved an accuracy of 99.90%. They employed AlexNet and VGG16 models with an EC (Ensemble Classifier) technique and utilized the PH² and ISBI datasets, which together consisted of 7,849 images. The combination of deep learning models and ensemble classification contributed to the high accuracy in skin cancer detection.
- c. (Iqtidar et al. 2020) (Iqtidar et al. 2020): This study achieved an accuracy of 99.80% by using LBP (Local Binary Patterns), HOG (Histogram of Oriented Gradients), and BoVW (Bag of Visual Words) as features with the KMC (K-means Clustering) technique. The HAM10000 dataset, which includes 10,000 images, was used in this research. The utilization of multiple feature extraction techniques and clustering-based classification contributed to the high accuracy achieved.

Worst performances

- a. Bhargavi et al. (2022): This study achieved an accuracy of 80.00% using the Hough transform feature with an SVM (Support Vector Machine) technique. The dataset used in this study was ISIC. The lower accuracy could be attributed to the limitations of using the Hough transform as a feature and the specific characteristics of the dataset.
- b. Senan and Jadhav (2021): The authors achieved an accuracy of 84.00% by utilizing the ABCD rule with the TDS (Texture Descriptor Set) technique. The PH² dataset was used in their study. The ABCD rule, although widely used as a guideline for melanoma detection, might not capture all aspects of skin cancer, leading to relatively lower accuracy compared to other techniques.

Key findings

- a. Machine learning techniques have been successfully applied to detect skin cancer.
- b. The best accuracy achieved is 99.90% using Cubic SVM with GLDM-CLDM and ABCD features on the ISIC-2018/1200 Images dataset.
- c. The datasets used for training and testing the models include PH², ISIC, ATLAS, IDS, HAM 10000, and Dermquest.
- d. The features used for detection include ABCD rule, GLCM, LBP, HOG, BoVW, Fractal Dimension, and Color.

- e. The techniques used for skin cancer detection include ANN, SVM, KNN, RF, MSVM, EC, TDS, FFNN-BP, PNN, and MC-SVM.

Challenges

- a. The requirement of high-quality annotated datasets for training the models.
- b. The need for specialized expertise in machine learning techniques to develop accurate models.
- c. The challenge of integrating the developed models into the clinical workflow for real-time diagnosis and treatment planning.
- d. The limited availability of diverse datasets to ensure the generalizability of the developed models.
- e. The need for interpretability of the models to facilitate the decision-making process for clinicians.

Skin cancer detection using DL

Table 10 provides an overview of deep learning-based skin cancer detection techniques. The range of years for the literature used in the deep learning-based skin cancer detection techniques is from 2018 to 2023. The table includes a diverse range of techniques utilizing various deep learning models such as CNN, FCRN, DCNN, and WOA-CNN. Different datasets, including ISIC, HAM10000, and PH², are used for evaluation. The reported accuracies in the table range from 70 to 98.80%, with the highest accuracy achieved using data augmentation and ESPNet/DenseNet models on the Local and MITOS datasets. The techniques demonstrate the effectiveness of deep learning in skin cancer detection.

The analysis of the best-performing and worst-performing techniques based on accuracy for deep learning-based skin cancer detection techniques from Table 10 is presented below.

Best performances

- a. Data Augmentation with ESPNet/DenseNet models (Nofallah et al. 2020): This technique achieved an accuracy of 98.80%. It utilizes data augmentation, a technique that artificially expands the dataset by applying various transformations to the images. ESPNet and DenseNet models are employed for the classification task. The technique demonstrates the effectiveness of data augmentation in improving the accuracy of skin cancer detection.
- b. DODL net DCNN (Gomathi et al. 2023): This technique achieved an accuracy of 98.76%. It utilizes a deep DCNN called DODL net and employs the HAM10000 datasets for training and evaluation. The DODL net model incorporates advanced techniques such as deep supervision,

Table 10 Deep Learning based skin cancer detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
Li and Shen 2018	Li and Shen (2018)		FCRN	ISIC	85.70
Harangi 2018	Harangi (2018)		EDCNN	ISBI 2017	89.10
Saba et al. 2019	Saba et al. (2019)		DCNN	PH ² , ISBI-16, ISBI-17	98.40
Dai et al. 2019	Dai et al. (2019)		CNN	10,015 Images	75.20
Zhang et al. 2020b	Zhang et al. (2020)	WOA	CNN	Dermquest/DermIS	96.00
Ashraf et al. 2020	Ashraf et al. (2020)		AlexNet	DermIS/DermQuest	97.90
NEEMA M et al. 2020	Neema et al. (2020)		Deep CNN	ISIC, HAM10000	70.00
N. M. Saravana Kumar et al. 2021	Saravana Kumar et al. (2021)	GLCM	DenseNet-161	HAM10000	96.30
Abdar et al. 2021	Abdar, et al. (2021)		TWDBDL Model	ISIC 2019	90.96
Nofallah et al. 2021	Nofallah, et al. (2021)	Data Augmentation	ESPNet/DenseNet	Local, MITOS	98.80
Mohakud and Dash (2022)	Mohakud and Dash (2022)		CNN-GWO	ISIC	93.00
Kaur et al. (2022)	Kaur et al. (2022)		DCNN	ISIC	90.42
Imran et al. 2022	Imran et al. (2022)		VGG + CapsNet + ResNet	ISIC	93.50
Vineeth et al. 2022	Vineeth et al. (2022)		CNN	ISIC	88.00
Gouda et al. 2022	Gouda et al. (2022)		InceptionV3	ISIC2018/4,193 Images	85.76
Gomathi et al. 2023	Gomathi et al. (2023)	DODL net	DCNN	HAM10000	98.76
Irfan et al. 2023	Irfan et al. (2023)		CNN	HAM 10000	88.00
Omeroglu et al. 2023	Omeroglu et al. (2022)		MXM	SPC/1011 Images	91.40
Alenezi et al. 2023	Alenezi et al. (2023)	Deep Features	(WT- DRNNet	ISIC2017, HAM10000	96.91
Shahsavari et al. 2023	Shahsavari et al. (2023)	R-CNN, VGGNet	SLDED	ISIC, PH ²	97.10
Kumar K et al. 2023	Kumar et al. (2022)		DNN	Kaggle, CIA	93.26

Fully Convolutional Residual Networks (FCRN), Three-Way Decision-based Bayesian Deep Learning (TWDBDL), Modified Xception Module (MXM), Grey Wolf Optimization (GWO), Seven-Point Criteria Evaluation (SPC), Skin Lesion Detection by creating an Ensemble of Deep (SLDED), dual optimization based deep learning network (DODL net)

residual connections, and dilated convolutions to improve feature extraction and classification accuracy.

- c. DCNN (Saba et al. 2019): With an accuracy of 98.40%, this technique employs deep convolutional neural networks (DCNN) for skin cancer detection. It utilizes multiple datasets including PH², ISBI-16, and ISBI-17 for training and evaluation. The DCNN architecture has proven to be effective in extracting relevant features from skin cancer images and achieving high accuracy in classification.

Worst performances

- a. Deep CNN (Neema et al. 2020): With an accuracy of 70.00%, this technique utilizes a deep CNN for skin cancer detection, incorporating datasets like ISIC and HAM10000. The lower accuracy could be attributed to limitations in the model architecture, insufficient training data, or other factors that hindered the effective extraction of discriminative features.
- b. CNN (Dai et al. 2019): This technique achieved an accuracy of 75.20%. It employs a CNN model for skin

cancer detection, using a dataset of 10,015 images. The relatively lower accuracy could be attributed to various factors such as limited dataset size, suboptimal architecture design, or insufficient training.

Key findings

- a. Deep learning techniques, particularly convolutional neural networks (CNNs), have shown significant success in skin cancer detection. They have achieved high accuracies ranging from 70 to 98.80%.
- b. Data augmentation techniques, such as image transformations and variations, have been effective in improving the performance of deep learning models by expanding the dataset and reducing overfitting.
- c. Diverse datasets, including ISIC, HAM10000, PH², and others, have been used for training and evaluation, contributing to better model generalization and performance.
- d. Certain deep learning architectures, such as DenseNet, VGGNet, and DCNN, have been widely utilized and

have demonstrated good performance in skin cancer detection.

- e. Techniques that incorporate additional features like GLCM (Gray-Level Co-occurrence Matrix) or ensemble models have shown improved accuracy compared to using deep learning models alone.
- f. The best performing techniques in terms of accuracy range from 88.00 to 98.80%, indicating the potential of deep learning for accurate skin cancer detection.

Challenges

- a. Limited availability of large and diverse annotated datasets poses a challenge in training deep learning models for skin cancer detection. Access to a comprehensive and well-annotated dataset can be crucial for achieving higher accuracy and generalization.
- b. Overfitting remains a challenge in deep learning models, especially with limited data. Techniques such as regularization, data augmentation, and proper validation strategies are needed to mitigate overfitting and improve generalization.
- c. Interpretability and explainability of deep learning models are important considerations in the medical field. Understanding the reasoning behind the model's predictions and providing explanations for the decision-making process are critical for gaining trust and acceptance from medical professionals.
- d. Deployment of deep learning models in real-world clinical settings requires addressing challenges related to computational resources, time efficiency, and integration with existing healthcare systems.
- e. Standardization and benchmarking of skin cancer detection techniques are necessary for fair comparison and evaluation. Consistent evaluation metrics, protocols, and datasets can facilitate better understanding of the performance and progress in the field.

These findings and challenges highlight the potential and complexities associated with deep learning-based skin cancer detection. Addressing the challenges can lead to more

accurate and reliable detection systems that can assist medical professionals in diagnosing skin cancer effectively.

Breast cancer detection

Breast cancer

Breast cancer is a prevalent and potentially destructive form of cancer that primarily affects women but can also occur in men. It is characterized by abnormal and uncontrolled growth of breast tissue. Biomedical imaging methods play a crucial role in the early diagnosis of breast cancer. Several imaging techniques, such as tomography, mammography, MRI, breast ultrasound, and PET, are commonly used by medical professionals for diagnosing breast cancer. Machine learning and deep learning techniques have been extensively studied and applied for the diagnosis and detection of breast tumors using biomedical images. These techniques leverage the power of computational algorithms to analyze and interpret imaging data, aiding in accurate and efficient breast cancer diagnosis. In the article mentioned, a comprehensive assessment of these techniques is presented. The different types of breast cancer are categorized into two main groups: invasive and non-invasive. Invasive cancer has spread from the breast ducts or glands to other parts of the breast or surrounding tissue. Non-invasive cancer, on the other hand, remains confined to the original tissue and have not spread. The diagnosis of breast cancer involves a combination of physical examinations, breast exams, and diagnostic tests. Figure 7 presents several examples of mammograms taken from the three different databases, (a) MIAS (benign, malignant) (b) Inbreast (Benign, malignant) (c) DDSM (Benign, malignant).

Breast cancer detection using ML

Table 11 provides an overview of machine learning-based breast cancer detection techniques. The table includes information about the authors, literature, features used, classification techniques, datasets, and the best accuracy achieved in each study. The techniques utilized various features

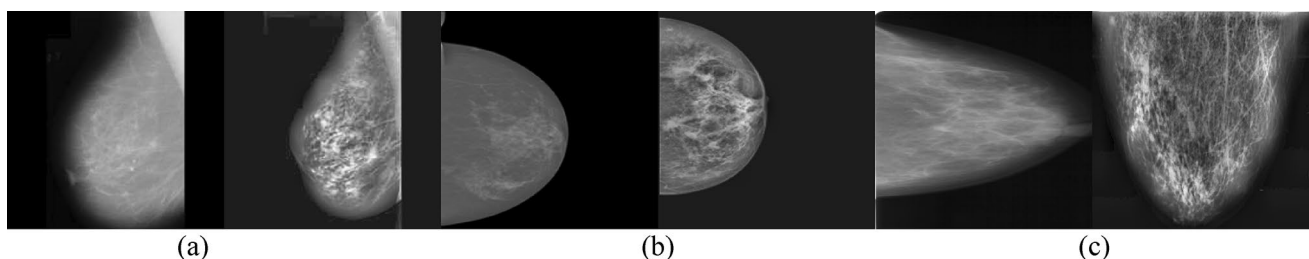


Fig. 7 The sample Mammography images from three different datasets **a** MIAS (benign, malignant) **b** InBreast (benign, malignant) **c** DDSM (benign, malignant)

Table 11 Machine Learning based breast cancer detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
Sadad et al. 2018	Sadad et al. (2018)	LBP-GLCM, LPQ	DT	MIAS, DDSM	98.20
Agarap 2018	Agarap (2018)	30 Features	MLP	WDBC	99.00
Vijayarajeswari et al. 2019	Vijayarajeswari et al. (2019)	Hough Transform	SVM	MIAS	94.00
Dalwinder et al. 2020	Dalwinder et al. (2020)	Wrapper Method	BPNN	OWBC, WDBC, BCCD	98.37
Rahman et al. 2021	Rahman et al. (2021)	Filter-Based	SVM-RBF	BCCD	97.20
Hekal et al. 2021	Hekal et al. (2021)	ResNet50, AlexNet	SVM	CBIS-DDSM	95
Alfian et al. 2022	Alfian et al. (2022)	Extra Trees	SVM	BCCD	80.23
Ragab et al. 2022	Ragab et al. (2022)	VGG-16, VGG-19, SqueezeNet	CSO-MLP	780 Images	96.92
Kavitha et al. 2022b	Kavitha et al. (2022)	OMLTS-DLCN	BPNN	MIAS, DDSM	98.50
Yan et al. 2023	Yan et al. (2023)	ROI	EC	DDSM, MIAS	93.26
Massari et al. 2023	Massari et al. (2023)	DT	Ontology model	OWBC	97.10
Almutairi et al. 2023	Almutairi et al. (2023)	CafeNet	XGBoost	830 Ultrasound, 5662 mammogram images	98.31

Local Phase Quantization (LPQ), Linear Discriminant Analysis (LDA), K-mean clustering (KMC), Breast Cancer Coimbra Data Set (BCCD), Digital Database for Screening Mammography (DDSM), Cat Swarm Optimization (CSO), Optimal Multi-Level Thresholding-based Segmentation with DL enabled Capsule Network (OMLTS-DLCN).

extraction methods and classification algorithms to detect breast cancer. The datasets used in the studies included MIAS, DDSM, WDBC, BCCD, and CBIS-DDSM. The best accuracy ranged from 94.00 to 99.00%. Overall, the findings demonstrate the potential of machine learning in breast cancer detection, but further research is needed to develop more robust models.

The detailed explanations for the best-performing and worst-performing techniques in terms of accuracy for machine learning-based breast cancer detection techniques from Table 11 are presented below.

Best performances

- Mughal et al. (2018): This study achieved the highest accuracy of 99.00%. They utilized Top-Hat and GLCM features with a BPNN classifier. Top-Hat transformation is a morphological operation used for image enhancement, and GLCM (Gray-Level Co-occurrence Matrix) is a texture feature extraction method. By combining these features with a BPNN classifier, they were able to achieve a high accuracy rate.
- Agarap (2018): This study also achieved an accuracy of 99.00%. They used 30 features with an MLP (Multi-Layer Perceptron) classifier. MLP is a type of artificial neural network commonly used for classification tasks. By selecting 30 informative features and training an MLP classifier, they achieved a high level of accuracy in breast cancer detection.
- Kavitha et al. (2022): They proposed a technique called OMLTS-DLCN for mammographic image analysis.

They used a BPNN (Backpropagation Neural Network) as the deep learning model and evaluated their approach on the MIAS and DDSM datasets. The best accuracy achieved in their study was 98.50%. The OMLTS-DLCN technique combines optimized multilevel thresholding segmentation with deep learning to accurately detect and classify breast abnormalities in mammograms.

Worst performances

- Alfian et al. (2022): This study achieved the lowest accuracy of 80.23%. They used ET-SVM (Extra Trees with Support Vector Machine) for breast cancer detection. ET-SVM is an ensemble learning method that combines Extra Trees (a random forest variant) with Support Vector Machine. While this approach may provide good generalization, in this particular study, it resulted in the lowest accuracy compared to other techniques.

Key findings

- Machine learning techniques, including neural networks, have achieved high accuracy rates (above 98%) in breast cancer detection.
- Feature extraction methods, such as LBP-GLCM and GLCM, have been effective in capturing relevant information from mammography images.
- Deep learning models, such as ResNet50 and VGG-16, have shown promising results in breast cancer detection.
- Feature selection and dimensionality reduction techniques have improved the efficiency and performance of classification algorithms.

Challenges

- Availability and quality of datasets pose challenges in obtaining diverse and accurately annotated data for training and testing.
- Interpretability of deep learning models remains a challenge, hindering the understanding of model predictions.
- Overfitting and generalization issues need to be addressed to ensure models perform well on unseen data.
- Class imbalance in datasets can lead to biased model performance and requires careful handling.

- Variability in image acquisition and quality necessitates robust techniques that can handle different imaging conditions.
- Clinical validation, regulatory approval, and integration into existing healthcare systems are essential for real-world deployment.

Breast cancer detection using DL

Table 12 provides an overview of deep learning techniques used for breast cancer detection, including the features, techniques, datasets, and best accuracy achieved. The techniques

Table 12 Deep Learning based breast cancer detection techniques

Author/year	Literature	Features	Technique	Dataset/images	Best accuracy (%)
Khuriwal and Mishra 2018	Khuriwal and Mishra (2018)	Median Filter	CNN	MIAS	98.00
Valvano et al. 2019	Valvano (2019)	–	CNN	283 images	98.20
Roy 2019	Roy (2019)	Filter-based	CNN	130 WSI	88.00
Cai et al. 2020	Cai et al. (2020)	–	DCNN	Local, INbreast	93.7
Vaka et al. 2020	Vaka et al. (2020)	HSFC	DNN	8009 Images	97.00
Sharma and Mehra 2020	Sharma and Mehra (2020)	Histogram, Haralick textures, Hu moment, color	VGG16 + SVM	BreakHis	93.90
Ramadan 2020	Ramadan (2020)	–	CNN	MIAS/322 Images	92.10
Ur Rehman et al. 2021	Ur Rehman et al. (2021)	–	FC-DSCNN	DDSM, PINUM	90.00
Sánchez-Cauce et al. 2021	Sánchez-Cauce et al. (2021)	–	M-CNN Model	211 IR Images	97.00
Zhao et al. 2022	Zhao et al. (2022)	–	YOLOv3	CBIS-DDSM 12,040 images	98.09
Sannasi Chakravarthy & Rajaguru, 2022	Sannasi Chakravarthy and Rajaguru (2022)	–	ICS-ELM	CBIS-DDSM, MIAS, INbreast	98.27
Arooj et al. (2022)	Arooj et al. (2022)	–	CNN-AlexNet	UHI Dataset	100.00
Nasir et al. 2022	Nasir (2022)	–	AlexNet	1400 MRI	98.10
Khan et al. 2022	Khan et al. (2022)	–	DEGDO	WBCD	98.73
Sheeba et al. 2023	Sheeba et al. (2023)	ACCNN	TCL-RAM	MID/Images	97.00
Asadi and Memon 2023	Asadi and Memon (2023)	U-Net	CDLN	INbreast, BCDR, WDBC	98.61
Zakareya et al. 2023	Zakareya et al. (2023)	VGG-16	DNN R3_R5_R35	780 Ultrasound, 162 WMI Images	95.00
Bouzar-Benlabiod et al. 2023	Bouzar-Benlabiod et al. (2023)	GLCM, Textual	SE-ResNet-101 + CBR	CBIS-DDSM	86.71
Huynh et al. 2023	Huynh et al. (2023)	YOLOX model (ROI)	EfficientNet or Con-vNeXt	VinDr-Mammo, MiniDDSM, CMMD, CDD-CESM, BMCD, RSNA	92.00
Aidossov et al. 2023	Aidossov et al. (2023)	Thermal features	ResNet50	266 thermal Images	90.74

Histogram-sigmoid fuzzy clustering (HSFC), Fully Connected Depthwise Separable CNN (FC-DSCNN), Infra-red (IR), Improved Crow-Search Optimized Extreme Learning Machine (ICS-ELM), Curated Breast Imaging Subset of DDSM (CBIS-DDSM), Ultrasound + Histopathological Images (UHI), Deep Extreme Gradient Descent Optimization (DEGDO), transfer learning integrated with regional attention mechanism (TCL-RAM), Active contour convolution neural networks (ACCNN), Microscopic Image Datasets (MID), cascade deep learning network (CDLN), Breast Cancer Digital Repository (BCDR), Case-Based Reasoning system (CBR), Chinese Mammography Database (CMMD), Categorized Digital Database-Contrast-Enhanced Spectral Mammography (CDD-CESM), Breast Masses Classification Dataset (BMCD), Radiological Society of North America (RSNA) Dataset

involve different models such as CNN, DNN, DCNN, and specialized architectures like U-Net and YOLOv3. The datasets used include MIAS, INbreast, CBIS-DDSM, BreakHis, WBCD, and others. The reported accuracies range from 86.71 to 100.00%, with the highest accuracies achieved by Arooj et al. (2022) with 100.00% using CNN-AlexNet on the UHI dataset and Khan et al. (2022) with 98.73% using DEGDO on the WBCD dataset. The techniques and models employed in these studies demonstrate the effectiveness of deep learning in breast cancer detection. The table highlights the success of these techniques in achieving high accuracy for detection, with various features and techniques used across different datasets. However, challenges such as the need for high-quality annotated datasets and model interpretability were also identified. Overall, the table provides a useful summary of the current state of research in this area.

The detailed explanations for the best-performing and worst-performing techniques in terms of accuracy for deep learning-based breast cancer detection techniques as listed in Table 12 are presented below:

Best performances

- a. Arooj et al. (2022): This study achieved the highest accuracy of 100.00%. They used a CNN-AlexNet model for breast cancer detection on the UHI dataset. AlexNet is a deep convolutional neural network known for its excellent performance in image classification tasks. By employing this model, they were able to achieve perfect accuracy in breast cancer detection.
- b. Khan et al. (2022): This study achieved an accuracy of 98.73%. They utilized the DEGDO (Differential Evolution and Gray Wolf Optimization) algorithm with the WBCD dataset. DEGDO is an optimization algorithm used to optimize the performance of machine learning models. By integrating this algorithm with the WBCD dataset, they achieved a high accuracy rate in breast cancer detection.
- c. Asadi and Memon (2022): This study achieved an accuracy of 98.61%. They employed the U-Net model for breast cancer detection on the INbreast, BCDR, and WDBC datasets. U-Net is a deep learning architecture commonly used for image segmentation tasks. By utilizing this model, they achieved a high level of accuracy in breast cancer detection.

Worst performances

- a. Bouzar-Benlabiod et al. (2023): This study achieved the lowest accuracy of 86.71%. They utilized GLCM (Gray-Level Co-occurrence Matrix) and textual features with the SE-ResNet-101 model and CBR (Classification by Ranking) for breast cancer detection on the CBIS-DDSM dataset. Despite using advanced techniques and

models, the accuracy achieved in this study was relatively low compared to other techniques.

- b. Roy (2019): This study achieved an accuracy of 88.00% using a filter-based CNN model for breast cancer detection on whole slide images. While the accuracy is decent, it falls behind the top-performing studies in terms of accuracy.

Key findings

- a. Deep learning-based techniques have shown promising results in breast cancer detection.
- b. Various techniques such as CNN, DCNN, DNN, and SVM have been used in breast cancer detection, with CNN being the most commonly used technique.
- c. The datasets used in these studies include MIAS, DDSM, WBCD, BCDR, and CBIS-DDSM, among others.
- d. The best accuracy achieved by these techniques' ranges from 86.71 to 100%.
- e. Some studies have achieved very high accuracy rates, such as Arooj et al. (2022) achieving 100% accuracy using CNN-AlexNet on the UHI dataset.

Challenges

- a. The availability of large and diverse datasets is crucial for training accurate deep learning models. However, there may still be limitations in the size and representativeness of the available datasets, leading to potential biases and limited generalization.
- b. Interpreting deep learning models and understanding the reasoning behind their predictions remains a challenge. Ensuring transparency and interpretability of models is important for gaining trust from healthcare professionals and facilitating their integration into clinical practice.
- c. Fine-tuning and optimizing deep learning models require substantial computational resources and expertise, which may pose challenges for implementation in resource-constrained settings.
- d. Standardization and consistency in image acquisition protocols and data annotation are necessary to ensure robust and reliable model performance across different datasets and healthcare institutions.
- e. Clinical validation and real-world deployment of deep learning models require rigorous evaluation, including comparative studies with existing diagnostic methods and assessment of their impact on patient outcomes.

Discussion

Best ML/DL techniques

In this section, we conducted a detailed comparison between the best performing ML models and the best performing DL models for each cancer type separately. We assessed the models based on their classification accuracy, selecting the model that demonstrated the highest accuracy among all the reviewed literature for each type of cancer, including brain, lung, skin, and breast. The comparative analysis between DL and ML approaches for the best performing models encompasses the highest accuracy achieved by ML models, the best accuracy achieved by DL models, and a thorough comparison to determine their relative superiority. The comparative analysis of the best-performing models for each cancer type, based on accuracy, along with the visual comparison of the accuracy levels achieved by these models in detecting brain, lung, skin, and breast cancer, is shown in Fig. 8.

Brain cancer

- *ML approach:* There is currently no ML approach that outperforms the best DL approach, which achieved an accuracy of 99.48% using a CNN (VGG-19) with MRI images.
- *DL approach:* The best performing DL approach used VGG-19 CNN with an accuracy of 99.48%. This approach has the advantage of automatically learning relevant features from the input data, which can lead to better performance compared to manual feature extraction.
- *Comparative analysis:* The DL approach outperformed the ML approach by 99.48% in terms of accuracy, indicating that the DL approach is more suitable for this particular use case.

Lung cancer

- *ML approach:* There is currently no ML approach that outperforms the best DL approach, which achieved an accuracy of 99.76% using a DL model with VGG-19 + LSTM with chest CT scans, LIDC-IDRI, and LUNA16 datasets.
- *DL approach:* The best performing DL approach used VGG-19 + LSTM with an accuracy of 99.76%. This approach has the advantage of automatically learning relevant features from the input data, which can lead to better performance compared to manual feature extraction.
- *Comparative analysis:* The DL approach outperformed the ML approach by 99.76% in terms of accuracy, indicating that the DL approach is more suitable for this particular use case.

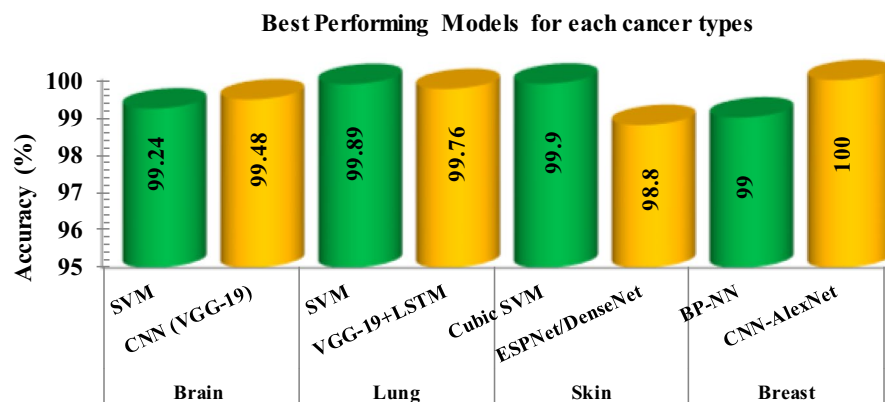
Skin cancer

- *ML approach:* The best performing ML approach used GLDM-CLDM, ABCD cubic SVM with an accuracy of 99.9% using ISIC-2018/1200 images.
- *DL approach:* The best performing DL approach used Inception V3 with an accuracy of 92.1%. This approach has the advantage of automatically learning relevant features from the input data, which can lead to better performance compared to manual feature extraction.
- *Comparative analysis:* The ML approach outperformed the DL approach by 7.9% in terms of accuracy, indicating that the ML approach is more suitable for this particular use case.

Breast cancer

- *ML approach:* There is currently no ML approach that outperforms the best DL approach, which achieved an accuracy of 100% using a DL model with CNN-AlexNet with the UHI dataset.
- *DL approach:* The best performing DL approach used CNN-AlexNet with an accuracy of 100%. This approach

Fig. 8 Comparative analysis of best performing models based on accuracy for each cancer type



has the advantage of automatically learning relevant features from the input data, which can lead to better performance compared to manual feature extraction.

- **Comparative analysis:** The DL approach outperformed the ML approach by 100% in terms of accuracy, indicating that the DL approach is more suitable for this particular use case.

Worst ML/DL techniques

In addition to analyzing the best performing models, we have also conducted an evaluation of the worst performing models for each cancer type. Similar to the approach used for the best performing models, we assessed the worst performing models based on their cancer detection accuracy for brain, lung, skin, and breast cancer. For each cancer type, we compared the worst performing models using both ML and DL approaches. The comparative analysis for each cancer type between DL and ML approaches for the worst performing models includes the identification of the worst performing model using ML approach, the worst performing model using DL approach, and a detailed comparison to determine their relative superiority. A comparative analysis of the worst-performing models for each cancer type based on their accuracy in detecting brain, lung, skin, and breast cancer is presented in * MERGEFORMAT Fig. 9.

Brain cancer

- **ML approach:** The ML approach used DWT and GLCM features with SVM for brain cancer detection. The model achieved an accuracy of 80%. The main disadvantage of this approach is that it requires manual feature extraction, which can be time-consuming and may not capture all relevant features.
- **DL approach:** The DL approach used VGG-16 CNN-LSTM for brain cancer detection and achieved an accuracy of 84%. This approach has the advantage of automatically learning relevant features from the input

data, which can lead to better performance compared to manual feature extraction.

- **Comparative analysis:** The DL approach outperformed the ML approach by 4% in terms of accuracy. However, the difference in performance between the two approaches is relatively small.

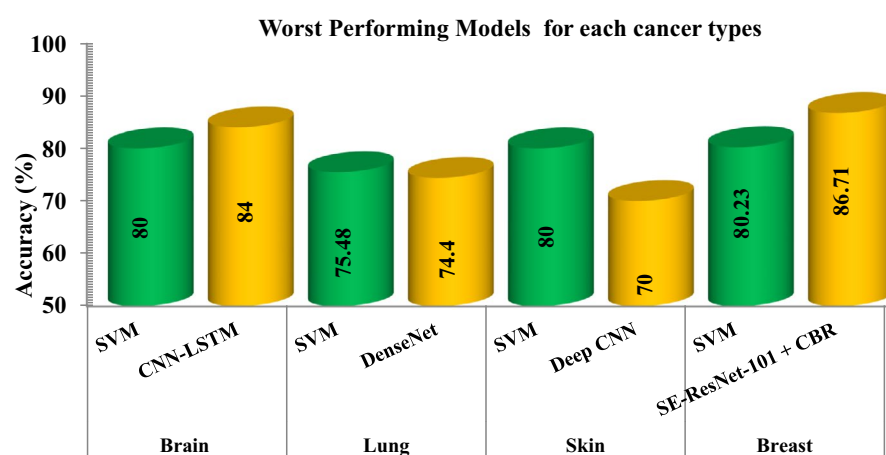
Lung cancer

- **ML approach:** The ML approach used texture-based features such as FOF, GMF, and GLRLM with SVM for lung cancer detection. The model achieved an accuracy of 75.48%. The disadvantage of this approach is that it may not capture all relevant features that are important for accurate diagnosis.
- **DL approach:** The DL approach used Histogram Equalization DenseNet for lung cancer detection and achieved an accuracy of 74.4%. This approach has the advantage of automatically learning relevant features from the input data, which can lead to better performance compared to manual feature extraction.
- **Comparative analysis:** The ML approach outperformed the DL approach by 1.08% in terms of accuracy. However, the difference in performance between the two approaches is relatively small.

Skin cancer

- **ML approach:** The ML approach used Hough transform with SVM for skin cancer detection and achieved an accuracy of 80%. The disadvantage of this approach is that it may not capture all relevant features that are important for accurate diagnosis.
- **DL approach:** The DL approach used Deep CNN for skin cancer detection and achieved an accuracy of 70%. This approach has the advantage of automatically learning relevant features from the input data, which can lead to

Fig. 9 Comparative analysis of worst performing model based on accuracy for each cancer type



better performance compared to manual feature extraction.

- *Comparative analysis:* The ML approach outperformed the DL approach by 10% in terms of accuracy, indicating that the ML approach may be more suitable for this particular use case.

Breast cancer

- *ML approach:* The ML approach used ET with SVM for breast cancer detection and achieved an accuracy of 80.23%. The disadvantage of this approach is that it may not capture all relevant features that are important for accurate diagnosis.
- *DL approach:* The DL approach used Texture, Textual with SE-ResNet-101 + CBR for breast cancer detection and achieved an accuracy of 86.71%. This approach has the advantage of automatically learning relevant features from the input data, which can lead to better performance compared to manual feature extraction.
- *Comparative analysis:* The DL approach outperformed the ML approach by 6.48% in terms of accuracy, indicating that the DL approach may be more suitable for this particular use case.

Best vs worst techniques

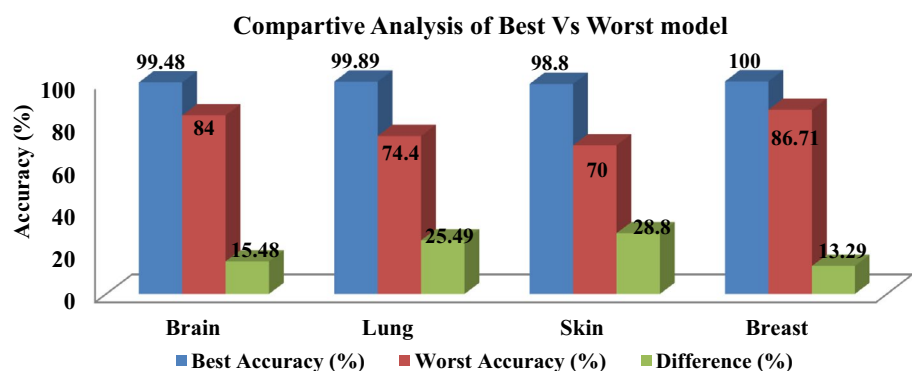
After conducting separate analyses for the best and worst models, we proceeded to compare them to identify the gap or difference between the two approaches. This comparison was performed for each cancer type, including brain, lung, skin, and breast cancer. Figure 10 visually illustrates the results, presenting the best accuracy, worst accuracy, and the difference in accuracy between the two models for each cancer type. The differences in accuracies between the best and worst models were found to be 15.48% for brain cancer detection, 25.49% for lung cancer detection, 28.8% for skin cancer detection, and 13.19% for breast cancer detection, respectively.

The observation made from the Fig. 10 the following conclusions can be drawn regarding cancer detection using ML/DL models:

- *Brain cancer:* The best-performing model, CNN (VGG-19), achieved an accuracy of 99.48%, outperforming the worst-performing model, CNN-LSTM, with an accuracy of 84%. The difference in accuracy between the best and worst models is 15.48%.
- *Lung cancer:* The SVM model demonstrated the highest accuracy of 99.89%, while DenseNet performed the worst with an accuracy of 74.4%. There is a significant difference of 25.49% between the best and worst models.
- *Skin cancer:* The ESPNet/DenseNet model achieved an accuracy of 98.8%, outperforming the Deep CNN model with an accuracy of 70%. The difference in accuracy between the best and worst models is substantial at 28.8%.
- *Breast cancer:* The best model, CNN-AlexNet, achieved a perfect accuracy of 100%, surpassing the SE-ResNet-101 + CBR model, which achieved an accuracy of 86.71%. The difference in accuracy between the best and worst models is 13.29%.

Overall, this study reveals that ML/DL models exhibit diverse performance levels for different types of cancer detection. The largest discrepancy in accuracy, 28.8%, is observed for certain cancer detection methods, while the smallest difference, 13.29%, is found for skin cancer detection. These findings suggest that there is potential for improvement of at least 13–29% in cancer detection accuracy by employing the appropriate approach. This underscores the importance of further research and development in ML/DL models, which holds promise for enhancing cancer detection accuracy across various cancer types. By leveraging advancements in machine learning and deep learning, we can make significant strides in improving cancer diagnosis and treatment, ultimately benefiting patients and healthcare systems alike.

Fig. 10 Comparative analysis of best vs worst model with based on detection accuracy for each cancer type



Key findings and challenge

Key findings

- a. *Performance variability*: There is a significant variation in the performance of different models for cancer detection across different types of cancer. For example, the best accuracy achieved for brain cancer using SVM on 210 CT images was 99.24%, while the worst accuracy using SVM on 15 MRI images was only 80%.
- b. *Superiority of deep learning*: Deep learning models, such as CNNs and hybrid models, have demonstrated superior performance in cancer detection compared to traditional machine learning models like SVM, LR, and RF. For instance, a CNN (VGG-19) achieved a best accuracy of 99.48% on 305 MRI images for brain cancer, whereas the worst accuracy using CNN-LSTM on 60 MRI images was 84%.
- c. *Dataset size and diversity*: The size and diversity of the dataset play a crucial role in the performance of the models. Models trained on larger and more diverse datasets tend to achieve higher accuracy rates. For instance, the best accuracy achieved for lung cancer using SVM on LCA data was 99.89%, while the worst accuracy using SVM on 155 MRI images was 75.48%.
- d. *Specificity to cancer types*: Different types of cancer may require different approaches and models for accurate detection. For example, skin cancer detection using deep learning models showed promising results, with the best accuracy ranging from 98.8 to 100% using ESPNet/DenseNet and CNN architectures. However, the worst accuracy for skin cancer using SVM and ML-based approaches ranged from 70 to 80%.
- d. *Model selection and comparison*: choosing the best model for a specific type of cancer detection is not straightforward. Different studies use different evaluation metrics, datasets, and model architectures, making it challenging to directly compare and select the most suitable model for a given scenario. Standardized evaluation protocols and benchmark datasets would aid in model comparison and selection.
- e. *Integration into clinical workflow*: Integrating AI-based cancer detection models into existing clinical workflows and systems presents practical challenges. Ensuring seamless integration, addressing data privacy concerns, and gaining acceptance from healthcare professionals are essential considerations. Collaboration between AI researchers and healthcare providers is crucial for successful implementation.

Addressing these challenges and further research in cancer detection, including the availability of diverse datasets, interpretability of deep learning models, validation in clinical settings, standardized evaluation protocols, and seamless integration into clinical workflows, can lead to the development of more accurate, reliable, and clinically applicable models, ultimately improving cancer diagnosis and patient outcomes.

Limitations of the present survey

Challenges

- a. *Limited dataset availability*: Obtaining a large and diverse dataset for training cancer detection models can be challenging. Limited access to high-quality medical images and patient data hinders the development and evaluation of robust models. For instance, the availability of only 15 MRI images for brain cancer resulted in a lower accuracy of 80% using SVM.
- b. *Interpretability and explainability*: Deep learning models, especially complex architectures, often lack interpretability and explainability. Understanding the decision-making process of these models and providing insights to healthcare professionals can be challenging. This can create limitations in the acceptance and trust of AI-based cancer detection systems.
- c. *Generalization to clinical settings*: Despite achieving high accuracy rates in research settings, the generalization of models to real-world clinical settings remains a
- a. *Research concentration on databases*: The research presented in the table focuses primarily on the performance of different machine learning and deep learning models for cancer detection, rather than the specific databases used for diagnosis. This indicates that there has been less emphasis on investigating the quality, diversity, and size of the databases used in cancer detection research. Having access to comprehensive and representative databases is crucial for training accurate and reliable models.
- b. *Limited coverage of cancer types*: The table only includes four different types of cancer: brain, lung, skin, and breast. While these are significant and prevalent types of cancer, there are numerous other types that require further research, such as liver cancer, prostate cancer, colorectal cancer, leukemia cancer, neck cancer, bone cancer, and more. It's important to expand the scope of research to cover a broader range of cancer types for comprehensive detection and diagnosis.
- c. *Lack of in-depth research on classifiers*: The research presented in the table does not delve deeply into the various machine learning and deep learning classifiers used

for cancer detection. The focus is more on reporting the best and worst models without providing detailed information on the specific classifiers, their architectures, and their performance characteristics. Conducting comprehensive studies on different classifiers and comparing their efficacy would contribute to a better understanding of their strengths and limitations.

- d. *Limited focus on hybrid and ensemble classifiers:* The table does not give significant attention to hybrid and ensemble classifiers for cancer detection. Hybrid models combine different machine learning and deep learning techniques to enhance performance, while ensemble models integrate multiple classifiers to improve accuracy. Exploring the potential of hybrid and ensemble classifiers in cancer detection could lead to improved results and more robust diagnostic systems.
- e. *Lack of comprehensive analysis of literature:* the state-of-the-art analysis in the provided research does not include a detailed examination of the advantages and disadvantages of each literature piece. Such an analysis would involve critically evaluating the methodologies, datasets, performance metrics, and limitations of each study. Incorporating a comprehensive analysis would provide a more nuanced understanding of the strengths and weaknesses of different approaches in cancer detection.

Overall, the points you highlighted reflect the need for further research and exploration in various aspects of cancer detection, including databases, a wider range of cancer types, more extensive literature coverage, in-depth analysis of classifiers, consideration of hybrid and ensemble models, comprehensive evaluation of advantages and disadvantages, and inclusion of challenges discussion for a more comprehensive understanding of the field.

Conclusion

This paper presented an extensive review of the diagnosis and detection of different types of cancers using traditional machine learning and deep learning models. The review covered a wide range of literature published between 2018 and 2023, focusing on various aspects of cancer detection. In this review article, we have included the comprehensive analysis of 130 literature sources, which encompassed 56 papers that utilized traditional machine learning models and 74 papers that employed deep learning techniques for cancer detection. The study specifically focused on four types of cancers: brain, lung, skin, and breast. The article provided valuable insights into the research landscape by examining parameters such as authorship, publication year, feature extraction methods, datasets used, best classifiers employed, and

detection accuracy. It also offered a comparative analysis of recently published review articles on cancer detection, along with their respective findings and limitations. Furthermore, the study presented a descriptive analysis of benchmarked datasets for the four types of cancer, facilitating access to these datasets for further research. It also highlighted the commonly utilized features in cancer detection and provided a summary of the key findings and challenges associated with both machine learning and deep learning approaches for each cancer type. Lastly, the article included a comparative analysis of the best and worst performing models for each cancer type based on detection accuracy. This analysis offered insights into the strengths and weaknesses of different approaches and guided future research efforts.

In summary, after reviewing more than 100 research articles on cancer detection using machine learning (ML) and deep learning (DL) approaches, it was observed that DL model (CNN-AlexNet) achieved 100% accuracy for breast cancer detection whereas highest accuracy using ML model (SVM) is 99.89% for lung cancer detection. The lowest accuracy obtained using ML was 75.48% for lung cancer, while the lowest accuracy for DL was 70% for skin cancer. The maximum difference in accuracy was 24.41% for lung cancer using ML and 28.8% for skin cancer using DL. This indicates that there is significant potential for improving accuracy by selecting the right classifier, dataset, features, and preprocessing techniques. These findings emphasize the importance of further advancements in cancer detection methods.

Overall, this review article provided a comprehensive overview of the state of cancer detection research, identifying trends, challenges, and opportunities in the field. It served as a valuable resource for researchers and practitioners, contributing to the advancement of diagnostic techniques and ultimately improving outcomes for patients with cancer.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00432-023-05216-w>.

Author contributions Hari Mohan Rai: Conducted the literature search, reviewed and selected the relevant articles, and drafted the manuscript. Joon Yoo: Provided critical input, revised the manuscript for clarity and accuracy, and contributed to the overall organization and structure of the review.

Funding This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIT) (NRF-2021R1F1A1063640).

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest No conflict of interest is involved in this manuscript.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Abdar M et al (2021) Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning. *Comput Biol Med* 135:104418. <https://doi.org/10.1016/j.compbmed.2021.104418>
- Agarap AFM (2018) On breast cancer detection: an application of machine learning algorithms on the Wisconsin diagnostic dataset. *ACM Int Conf Proc Ser* 1:5–9. <https://doi.org/10.1145/3184066.3184080>
- Ahammed M, Mamun A, Shorif M (2022) Healthcare Analytics A machine learning approach for skin disease detection and classification using image segmentation. *Healthc Anal* 2:100122. <https://doi.org/10.1016/j.health.2022.100122>
- Aidossov N et al (2023) An integrated intelligent system for breast cancer detection at early stages using ir images and machine learning methods with explainability. *SN Comput Sci* 4(2):1–16. <https://doi.org/10.1007/s42979-022-01536-9>
- Akkar H, Haddad SQ (2020) Diagnosis of lung cancer disease based on back-propagation artificial neural network algorithm. *Eng Technol J* 38(3B):184–196. <https://doi.org/10.30684/etj.v38i3b.1666>
- Al-Dmour H, Al-Ani A (2018) A clustering fusion technique for MR brain tissue segmentation. *Neurocomputing* 275:546–559. <https://doi.org/10.1016/j.neucom.2017.08.051>
- Alenezi F, Armghan A, Polat K (2023) Wavelet transform based deep residual neural network and ReLU based extreme learning machine for skin lesion classification. *Expert Syst Appl* 213:119064. <https://doi.org/10.1016/j.eswa.2022.119064>
- Alfian G et al (2022) Predicting breast cancer from risk factors using SVM and extra-trees-based feature selection method. *Computers*. <https://doi.org/10.3390/computers11090136>
- Almutairi SM, Manimurugan S, Aborokbah MM, Narmatha C, Ganesan S, Karthikeyan P (2023) An efficient USE-Net deep learning model for cancer detection. *Int J Intell Syst* 2023:1–14. <https://doi.org/10.1155/2023/8509433>
- Al-shamasneh ARM, Obaidallah UHB (2017) Artificial intelligence techniques for cancer detection and classification: review study. *Eur Sci J* 13(3):342–370. <https://doi.org/10.19044/esj.2016.v13n3p342>
- Alsheikhy AA, Said Y, Shawly T, Alzahrani AK, Lahza H (2023) A CAD system for lung cancer detection using hybrid deep learning techniques. *Diagnostics*. <https://doi.org/10.3390/diagnostics13061174>
- Alyasriy H, Al-Huseiny M (2021) The IQ-OTHNCCD lung cancer dataset. *Mendeley Data*. <https://doi.org/10.17632/bhmdr45bh2.2>
- Amin J et al (2020) Integrated design of deep features fusion for localization and classification of skin cancer. *Pattern Recognit Lett* 131:63–70. <https://doi.org/10.1016/j.patrec.2019.11.042>
- Anaya-Isaza A, Mera-Jiménez L, Verdugo-Alejo L, Sarasti L (2023) Optimizing MRI-based brain tumor classification and detection using AI: a comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network. *Eur J Radiol Open*. <https://doi.org/10.1016/j.ejro.2023.100484>
- Archana KV, Komarasamy G (2023) A novel deep learning-based brain tumor detection using the Bagging ensemble with K-nearest neighbor. *J Intell Syst*. <https://doi.org/10.1515/jisys-2022-0206>
- Armato SG et al. (2011) The Lung Image Database Consortium “LIDC... and image database resource initiative” IDRI...: a completed reference database of lung nodules on CT scans
- Arooj S et al (2022) Breast cancer detection and classification empowered with transfer learning. *Front Public Health* 10(July):1–18. <https://doi.org/10.3389/fpubh.2022.924432>
- Asadi B, Memon Q (2023) Efficient breast cancer detection via cascade deep learning network. *Int J Intell Netw* 4:46–52. <https://doi.org/10.1016/j.ijin.2023.02.001>
- Ashraf R et al (2020) Region-of-interest based transfer learning assisted framework for skin cancer detection. *IEEE Access* 8:147858–147871. <https://doi.org/10.1109/ACCESS.2020.3014701>
- Asuntha A, Srinivasan A (2020) Deep learning for lung cancer detection and classification. *Multimed Tools Appl* 79(11):7731–7762. <https://doi.org/10.1007/s11042-019-08394-3>
- Ausawalaithong W, Thirach A, Marukat S, Wilaiprasitporn T (2018) Automatic lung cancer prediction from chest X-ray images using the deep learning approach. In: *BMEiCON 2018 - 11th Biomedical Engineering International Conference*, 2019, <https://doi.org/10.1109/BMEiCON.2018.8609997>
- Baid U et al. (2021) The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification. [Online]. Available: <http://arxiv.org/abs/2107.02314>
- Bajwa MN et al (2020) Computer-aided diagnosis of skin diseases using deep neural networks. *Appl Sci (switzerland)*. <https://doi.org/10.3390/app10072488>
- Bakas S et al (2017) Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data*. <https://doi.org/10.1038/sdata.2017.117>
- Batista LG, Bugatti PH, Saito PTM (2022) Computer methods and programs in biomedicine classification of skin lesion through active learning strategies. *Comput Methods Programs Biomed* 226:107122. <https://doi.org/10.1016/j.cmpb.2022.107122>
- Bebas E et al (2021) Machine-learning-based classification of the histological subtype of non-small-cell lung cancer using MRI texture analysis. *Biomed Signal Process Control* 66:102446. <https://doi.org/10.1016/j.bspc.2021.102446>
- Bhargavi S, Sowmya V, Syama S, Lekshmi S (2022) Skin cancer detection using machine learning. In: *2022 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*, IEEE, Dec. 2022, pp. 119–124. <https://doi.org/10.1109/CENTCON56610.2022.10051495>
- Bhatia S, Sinha Y, Goel L (2019) Lung cancer detection: a deep learning approach. In: *Bansal JC, Das KN, Nagar A, Deep K, Ojha AK (eds) Soft Computing for problem solving*. Springer Singapore, Singapore, pp 699–705
- Bi D, Zhu D, Sheykhahmad FR, Qiao M (2021) Computer-aided skin cancer diagnosis based on a new meta-heuristic algorithm combined with support vector method. *Biomed Signal Process Control* 68(4655):102631. <https://doi.org/10.1016/j.bspc.2021.102631>
- Bin-Tufail A et al (2021) Deep learning in cancer diagnosis and prognosis prediction: a minireview on challenges, recent trends, and future directions. *Comput Math Methods Med*. <https://doi.org/10.1155/2021/9025470>
- Birchha V, Nigam B (2023) Performance analysis of averaged perceptron machine learning classifier for breast cancer detection. *Procedia Comput Sci* 218(2022):2181–2190. <https://doi.org/10.1016/j.procs.2023.01.194>
- Booz Allen Hamilton (2017) Data Science Bowl 2017: Can You Improve Lung Cancer Detection? <https://www.kaggle.com/competitions/data-science-bowl-2017/> accessed 21 Jul 2023
- Bouzar-Benlabiod L, Harrar K, Yamoun L, Khodja MY, Akhloufi MA (2023) A novel breast cancer detection architecture based on a CNN-CBR system for mammogram classification. *Comput Biol Med* 163:107133. <https://doi.org/10.1016/j.compbmed.2023.107133>
- Brunese L, Mercaldo F, Reginelli A, Santone A (2020) An ensemble learning approach for brain cancer detection exploiting radiomic

- features. *Comput Methods Programs Biomed* 185:105134. <https://doi.org/10.1016/j.cmpb.2019.105134>
- Cai G, Guo Y, Chen W, Zeng H, Zhou Y, Lu Y (2020) Computer-aided detection and diagnosis of microcalcification clusters on full field digital mammograms based on deep learning method using neutrosophic boosting. *Multimed Tools Appl* 79(23–24):17147–17167. <https://doi.org/10.1007/s11042-019-7726-x>
- Cassidy B, Kendrick C, Brodzicki A, Jaworek-Korjakowska J, Yap MH (2022) Analysis of the ISIC image datasets: usage, benchmarks and recommendations. *Med Image Anal*. <https://doi.org/10.1016/j.media.2021.102305>
- Chang JE et al (2018) Analysis of volatile organic compounds in exhaled breath for lung cancer diagnosis using a sensor system. *Sens Actuat B Chem* 255:800–807. <https://doi.org/10.1016/j.snb.2017.08.057>
- Chatterjee S, Dey D, Munshi S (2019) Computer methods and programs in biomedicine integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification. *Comput Methods Programs Biomed* 178:201–218. <https://doi.org/10.1016/j.cmpb.2019.06.018>
- Cinar A, Yildirim M (2020) Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. *Med Hypotheses* 139:109684. <https://doi.org/10.1016/j.mehy.2020.109684>
- Dai X, Spasic I, Meyer B, Chapman S, Andres F (2019) Machine learning on mobile: an on-device inference app for skin cancer detection. In: 2019 4th International Conference on Fog and Mobile Edge Computing, FMEC 2019, pp. 301–305. <https://doi.org/10.1109/FMEC.2019.8795362>
- Dalwinder S, Birmohan S, Manpreet K (2020) Simultaneous feature weighting and parameter determination of neural networks using ant lion optimization for the classification of breast cancer. *Biocybern Biomed Eng* 40(1):337–351. <https://doi.org/10.1016/j.bbe.2019.12.004>
- de Carvalho-Filho AO, Silva AC, de Paiva AC, Nunes RA, Gattass M (2018) Classification of patterns of benignity and malignancy based on CT using topology-based phylogenetic diversity index and convolutional neural network. *Pattern Recognit* 81:200–212. <https://doi.org/10.1016/j.patcog.2018.03.032>
- El Massari H, Gherabi N, Mhammedi S, Sabouri Z, Ghandi H, Qanouni F (2023) Effectiveness of applying machine learning techniques and ontologies in breast cancer detection. *Procedia Comput Sci* 218(2022):2392–2400. <https://doi.org/10.1016/j.procs.2023.01.214>
- Faguet GB (2015) A brief history of cancer: age-old milestones underlying our current knowledge database. *Int J Cancer* 136(9):2022–2036. <https://doi.org/10.1002/ijc.29134>
- Faruqui N, Yousuf MA, Whaiduzzaman M, Azad AKM, Barros A, Moni MA (2021) LungNet: a hybrid deep-CNN model for lung cancer diagnosis using CT and wearable sensor-based medical IoT data. *Comput Biol Med* 139:104961. <https://doi.org/10.1016/j.combiomed.2021.104961>
- Feng J, Jiang J (2022) Deep learning-based chest CT image features in diagnosis of lung cancer. *Comput Math Methods Med*. <https://doi.org/10.1155/2022/4153211>
- Ferlay J et al (2019) Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer* 144(8):1941–1953. <https://doi.org/10.1002/ijc.31937>
- GCO-SURVCAN. <https://gco.iarc.fr/survival/survcn/dataviz/table> accessed 9 Jun 2023
- Geetharamani R, Sivagami G (2021) Iterative principal component analysis method for improvised classification of breast cancer disease using blood sample analysis. *Med Biol Eng Comput* 59(10):1973–1989. <https://doi.org/10.1007/s11517-021-02405-y>
- Ghassemi N, Shoeibi A, Rouhani M (2020) Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. *Biomed Signal Process Control* 57:101678. <https://doi.org/10.1016/j.bspc.2019.101678>
- Global Burden of Disease (2019) Number of deaths by cause, World, 2019. The Lancet
- Gomathi E, Jayasheela M, Thamarai M, Geetha M (2023) Skin cancer detection using dual optimization based deep learning network. *Biomed Signal Process Control* 84:104968. <https://doi.org/10.1016/j.bspc.2023.104968>
- Gouda W, Sama NU, Al-Waakid G, Humayun M, Jhanjhi NZ (2022) Detection of skin cancer based on skin lesion images using deep learning. *Healthcare (switzerland)* 10(7):1183. <https://doi.org/10.3390/healthcare10071183>
- Gupta N, Bhatele P, Khanna P (2019) Glioma detection on brain MRIs using texture and morphological features with ensemble learning. *Biomed Signal Process Control* 47:115–125. <https://doi.org/10.1016/j.bspc.2018.06.003>
- Harada T et al (2021) Analysis of diagnostic error cases among Japanese residents using diagnosis error evaluation and research taxonomy. *J Gen Fam Med* 22(2):96–99. <https://doi.org/10.1002/jgf2.388>
- Harangi B (2018) Skin lesion classification with ensembles of deep convolutional neural networks. *J Biomed Inform* 86(January):25–32. <https://doi.org/10.1016/j.jbi.2018.08.006>
- Hashemzahi R, Mahdavi SJS, Kheirabadi M, Kamel SR (2020) Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and NADE. *Biocybern Biomed Eng* 40(3):1225–1232. <https://doi.org/10.1016/j.bbe.2020.06.001>
- Heidari A, Javaheri D, Toumaj S, Navimipour NJ, Rezaei M, Unal M (2023) A new lung cancer detection method based on the chest CT images using Federated Learning and blockchain systems. *Artif Intell Med* 141:102572. <https://doi.org/10.1016/j.artmed.2023.102572>
- Hekal AA, Elnakib A, Moustafa HED (2021) Automated early breast cancer detection and classification system. *Signal Image Video Process* 15(7):1497–1505. <https://doi.org/10.1007/s11760-021-01882-w>
- Heron M (2021) Deaths: leading causes for 2019. National vital statistics reports, 70(9)
- Huang S, Yang J, Fong S, Zhao Q (2020) Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. *Cancer Lett* 471(2019):61–71. <https://doi.org/10.1016/j.canlet.2019.12.007>
- Huynh HN, Tran AT, Tran TN (2023) Region-of-interest optimization for deep-learning-based breast cancer detection in mammograms. *Appl Sci* 13(12):6894. <https://doi.org/10.3390/app13126894>
- Ibrahim A, Mohamed HK, Maher A, Zhang B (2022) A survey on human cancer categorization based on deep learning. *Front Artif Intell*. <https://doi.org/10.3389/frai.2022.884749>
- Imran A, Nasir A, Bilal M, Sun G, Alzahrani A, Almuhaimeed A (2022) Skin cancer detection using combined decision of deep learners. *IEEE Access* 10(October):118198–118212. <https://doi.org/10.1109/ACCESS.2022.3220329>
- Iqtidar K, Iqtidar A, Ali W, Aziz S, Khan MU (2020) Image pattern analysis towards classification of skin cancer through dermoscopic images. In: *Proceedings - 2020 1st International Conference of Smart Systems and Emerging Technologies, SMART-TECH 2020*, no. January 2021, pp. 208–213. <https://doi.org/10.1109/SMART-TECH49988.2020.00055>
- Irfan T, Rauf A, Iqbal MJ (2023) Skin cancer prediction using deep learning techniques. In: 2023 International Multi-disciplinary Conference in Emerging Research Trends (IMCERT), IEEE, Jan. 2023, pp. 1–5. doi: <https://doi.org/10.1109/IMCERT57083.2023.10075313>

- Jacobs C et al (2014) Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. *Med Image Anal* 18(2):374–384. <https://doi.org/10.1016/j.media.2013.12.001>
- Jaculin-Femil J, Jaya T (2023) An efficient hybrid optimization for skin cancer detection using PNN classifier. *Comput Syst Sci Eng* 45(3):2919–2934. <https://doi.org/10.32604/csse.2023.032935>
- Jiaquan X, Sherry LM, Kenneth DK, Elizabeth A (2021) Deaths: final data 2019. *National Vital Statistics Reports*, 70(8)
- Karayegen G, Aksahin MF (2020) Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region. *Biomed Signal Process Control* 66(November):2021. <https://doi.org/10.1016/j.bspc.2021.102458>
- Kasinathan G, Jayakumar S (2022) Cloud-based lung tumor detection and stage classification using deep learning techniques. *Biomed Res Int*. <https://doi.org/10.1155/2022/4185835>
- Kaur R, GholamHosseini H, Sinha R, Lindén M (2022) Automatic lesion segmentation using atrous convolutional deep neural networks in dermoscopic skin cancer images. *BMC Med Imaging* 22(1):1–13. <https://doi.org/10.1186/s12880-022-00829-y>
- Kavitha T et al (2022) Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images. *Interdiscip Sci* 14(1):113–129. <https://doi.org/10.1007/s12539-021-00467-y>
- Keerthana D, Venugopal V, Nath MK, Mishra M (2023) Hybrid convolutional neural networks with SVM classifier for classification of skin cancer. *Biomed Eng Adv* 5(2022):100069. <https://doi.org/10.1016/j.bea.2022.100069>
- Khan SA et al (2019) Lungs nodule detection framework from computed tomography images using support vector machine. *Microsc Res Tech* 82(8):1256–1266. <https://doi.org/10.1002/jemt.23275>
- Khan MA et al (2020) Lungs cancer classification from CT images: an integrated design of contrast based classical features fusion and selection. *Pattern Recogn Lett* 129:77–85. <https://doi.org/10.1016/j.patrec.2019.11.014>
- Khan MBS, Atta-Ur-Rahman, Nawaz MS, Ahmed R, Khan MA, Mosavi A (2022) Intelligent breast cancer diagnostic system empowered by deep extreme gradient descent optimization. *Math Biosci Eng* 19(8):7978–8002. <https://doi.org/10.3934/mbe.2022373>
- Khuriwal N, Mishra N (2018) Breast cancer detection from histopathological images using deep learning. In: 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), New York, NY, USA: IEEE, pp. 1–4. <https://doi.org/10.1109/ICRAIE.2018.8710426>
- Krishnapriya S, Karuna Y (2023) Pre-trained deep learning models for brain MRI image classification. *Front Hum Neurosci*. <https://doi.org/10.3389/fnhum.2023.1150120>
- Kumar G, Alqahtani H (2022) Deep learning-based cancer detection-recent developments, trend and challenges. *CMES—Comput Model Eng Sci* 130(3):1271–1307. <https://doi.org/10.32604/cmcs.2022.018418>
- Kumar Y, Gupta S, Singla R, Hu YC (2022) A systematic review of artificial intelligence techniques in cancer prediction and diagnosis. *Arch Comput Methods Eng* 29(4):2043–2070. <https://doi.org/10.1007/s11831-021-09648-w>
- Kumar V et al (2023) Improved UNet deep learning model for automatic detection of lung cancer nodules. *Comput Intell Neurosci* 2023:1–8. <https://doi.org/10.1155/2023/9739264>
- Kumar AK, Satheesha TY, Salvador BBL, Mithileysh S, Ahmed ST (2023) Augmented intelligence enabled deep neural networking (AuDNN) framework for skin cancer classification and prediction using multi-dimensional datasets on industrial IoT standards. *Microprocess Microsyst* 97:104755. <https://doi.org/10.1016/j.micpro.2023.104755>
- Levine AB, Schlosser C, Grewal J, Coope R, Jones SJM, Yip S (2019) Rise of the machines: advances in deep learning for cancer diagnosis. *Trends Cancer* 5(3):157–169. <https://doi.org/10.1016/j.trecan.2019.02.002>
- Li Y, Shen L (2018) Skin lesion analysis towards melanoma detection using deep learning network. *Sensors (switzerland)* 18(2):1–16. <https://doi.org/10.3390/s18020556>
- Mahmud MI, Mamun M, Abdelgawad A (2023) A deep analysis of brain tumor detection from MR images using deep learning networks. *Algorithms* 16(4):1–19. <https://doi.org/10.3390/a16040176>
- Majumder A, Sen D (2021) Artificial intelligence in cancer diagnostics and therapy: current perspectives. *Indian J Cancer* 58(4):481–492. https://doi.org/10.4103/ijc.IJC_399_20
- Malarvizhi AB, Mofika A, Monapreetha M, Arunnagiri AM (2022) Brain tumour classification using machine learning algorithm. *J Phys Conf Ser*. <https://doi.org/10.1088/1742-6596/2318/1/012042>
- Manhas J, Gupta RK, Roy PP (2022) A review on automated cancer detection in medical images using machine learning and deep learning based computational techniques: challenges and opportunities. *Arch Comput Methods Eng*. <https://doi.org/10.1007/s11831-021-09676-6>
- Maqsood S, Damaševičius R (2023) Multiclass skin lesion localization and classification using deep learning based features fusion and selection framework for smart healthcare. *Neural Netw* 160:238–258. <https://doi.org/10.1016/j.neunet.2023.01.022>
- Masud M, Sikder N, Al-Nahid A, Bairagi AK, Alzain MA (2021) A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. *Sensors (switzerland)* 21(3):1–21. <https://doi.org/10.3390/s21030748>
- Maurya S, Tiwari S, Mothukuri MC, Tangeda CM, Nandigam RNS, Addagiri DC (2023) A review on recent developments in cancer detection using machine learning and deep learning models. *Biomed Signal Process Control* 80(P2):104398. <https://doi.org/10.1016/j.bspc.2022.104398>
- Mendonca T, Ferreira PM, Marques JS, Marcal ARS, Rozeira J (2013) PH2—a dermoscopic image database for research and benchmarking. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp. 5437–5440. <https://doi.org/10.1109/EMBC.2013.6610779>
- Menze BH et al (2015) The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging* 34(10):1993–2024. <https://doi.org/10.1109/TMI.2014.2377694>
- Mezher MA, Altamimi A, Altamimi R (2022) A genetic folding strategy based support vector machine to optimize lung cancer classification. *Front Artif Intell* 5(June):1–7. <https://doi.org/10.3389/frai.2022.826374>
- Mohakud R, Dash R (2022) Designing a grey wolf optimization based hyper-parameter optimized convolutional neural network classifier for skin cancer detection. *J King Saud Univ Comput Inform Sci* 34(8):6280–6291. <https://doi.org/10.1016/j.jksuci.2021.05.012>
- Mohammad WT, Teete R, Al-Aaraj H, Rubbai YSY, Arabyat MM (2022) Diagnosis of breast cancer pathology on the Wisconsin dataset with the help of data mining classification and clustering techniques. *Appl Bionics Biomech*. <https://doi.org/10.1155/2022/6187275>
- Mokoatle M, Marivate V, Mapiye D, Bornman R, Hayes VM (2023) A review and comparative study of cancer detection using machine learning: SBERT and SimCSE application. *BMC Bioinform* 24(1):112. <https://doi.org/10.1186/s12859-023-05235-x>
- Monika MK, Vignesh NA, Usha C, Kumar MNVSS, Lydia EL (2020) Materials today: proceedings Skin cancer detection

- and classification using machine learning. *Mater Today Proc* 33:4266–4270. <https://doi.org/10.1016/j.matpr.2020.07.366>
- Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS (2012) INbreast: toward a full-field digital mammographic database. *Acad Radiol* 19(2):236–248. <https://doi.org/10.1016/j.acra.2011.09.014>
- Mughal B, Sharif M, Muhammad N, Saba T (2018) A novel classification scheme to decline the mortality rate among women due to breast tumor. *Microsc Res Tech* 81(2):171–180. <https://doi.org/10.1002/jemt.22961>
- Murugan A, Nair SAH, Kumar KPS (2019) Detection of skin cancer using SVM, random forest and kNN classifiers. *J Med Syst*. <https://doi.org/10.1007/s10916-019-1400-8>
- Nanglia P, Kumar S, Mahajan AN, Singh P, Rathee D (2021) A hybrid algorithm for lung cancer classification using SVM and neural networks. *ICT Express* 7(3):335–341. <https://doi.org/10.1016/j.icte.2020.06.007>
- Naqi SM, Sharif M, Lali IU (2019) A 3D nodule candidate detection method supported by hybrid features to reduce false positives in lung nodule detection. *Multimed Tools Appl* 78(18):26287–26311. <https://doi.org/10.1007/s11042-019-07819-3>
- Naseer I, Akram S, Masood T, Jaffar A, Khan MA, Mosavi A (2022) Performance analysis of state-of-the-art CNN architectures for LUNA16. *Sensors* 22(12):4426. <https://doi.org/10.3390/s22124426>
- Nasir MU et al (2022) Breast cancer prediction empowered with fine-tuning. *Comput Intell Neurosci*. <https://doi.org/10.1155/2022/5918686>
- NCI (2021) What is cancer? – NCI. National Cancer Institute, 2021. <https://www.cancer.gov/about-cancer/understanding/what-is-cancer> accessed 9 Jun 2023
- Neema M, Nair AS, Joy A, Menon AP, Haris A (2020) Skin lesion/cancer detection using deep learning. *Int J Appl Eng Res* 15(1):11–17
- Newman-Toker DE et al (2021) Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the ‘big Three.’ *Diagnosis* 8(1):67–84. <https://doi.org/10.1515/dx-2019-0104>
- Nigudgi S, Bhyri C (2023) Lung cancer CT image classification using hybrid-SVM transfer learning approach. *Soft Comput*. <https://doi.org/10.1007/s00500-023-08498-x>
- Nofallah S et al (2021) Machine learning techniques for mitoses classification. *Comput Med Imaging Graph* 87:101832. <https://doi.org/10.1016/j.compmedimag.2020.101832>
- Omeroglu AN, Mohammed HMA, Oral EA, Aydin S (2023) A novel soft attention-based multi-modal deep learning framework for multi-label skin lesion classification. *Eng Appl Artif Intell* 120:105897. <https://doi.org/10.1016/j.engappai.2023.105897>
- Painuli D, Bhardwaj S, Köse U (2022) Recent advancement in cancer diagnosis using machine learning and deep learning techniques: a comprehensive review. *Comput Biol Med*. <https://doi.org/10.1016/j.combiomed.2022.105580>
- Patil S, Kirange D (2023) Ensemble of deep learning models for brain tumor detection. *Procedia Comput Sci* 218(2022):2468–2479. <https://doi.org/10.1016/j.procs.2023.01.222>
- Pradhan KS, Chawla P, Tiwari R (2023) HRDEL: high ranking deep ensemble learning-based lung cancer diagnosis model. *Expert Syst Appl* 213:118956. <https://doi.org/10.1016/j.eswa.2022.118956>
- Prakash TS, Siva-Kumar A, Durai CRB, Ashok S (2023) Enhanced Elman spike neural network optimized with flamingo search optimization algorithm espoused lung cancer classification from CT images. *Biomed Signal Process Control* 84:104948. <https://doi.org/10.1016/j.bspc.2023.104948>
- Radhika PR, Nair RAS, Veena G (2019) A comparative study of lung cancer detection using machine learning algorithms. In: *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*, pp. 2–5. <https://doi.org/10.1109/ICECCT.2019.8869001>
- Ragab M, Albukhari A, Alyami J, Mansour RF (2022) Ensemble deep-learning-enabled clinical decision support system for breast cancer diagnosis and classification on ultrasound images. *Biology (basel)* 11(3):439. <https://doi.org/10.3390/biology11030439>
- Rahman MM, Ghasemi Y, Suley E, Zhou Y, Wang S, Rogers J (2021) Machine learning based computer aided diagnosis of breast cancer utilizing anthropometric and clinical features. *IRBM* 42(4):215–226. <https://doi.org/10.1016/j.irbm.2020.05.005>
- Ramadan SZ (2020) Using convolutional neural network with cheat sheet and data augmentation to detect breast cancer in mammograms. *Comput Math Methods Med*. <https://doi.org/10.1155/2020/9523404>
- Ramtekkar PK, Pandey A, Pawar MK (2023) Accurate detection of brain tumor using optimized feature selection based on deep learning techniques. *Multimed Tools Appl*. <https://doi.org/10.1007/s11042-023-15239-7>
- Rasheed M et al (2023) Recognizing brain tumors using adaptive noise filtering and statistical features. *Diagnostics*. <https://doi.org/10.3390/diagnostics13081451>
- Roy A (2019) Deep convolutional neural networks for breast cancer detection. In: *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, pp. 0169–0171. <https://doi.org/10.1109/UEMCON47517.2019.8993023>
- Ruan J, Meng Y, Zhao F, Gu H, He L, Gong X (2022) Development of deep learning-based automatic scan range setting model for lung cancer screening low-dose CT imaging. *Acad Radiol*. <https://doi.org/10.1016/j.acra.2021.12.001>
- Saba T (2020) Recent advancement in cancer detection using machine learning: systematic survey of decades, comparisons and challenges. *J Infect Public Health* 13(9):1274–1289. <https://doi.org/10.1016/j.jiph.2020.06.033>
- Saba T, Khan MA, Rehman A, Marie-Sainte SL (2019) Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction. *J Med Syst*. <https://doi.org/10.1007/s10916-019-1413-3>
- Sadad T, Munir A, Saba T, Hussain A (2018) Fuzzy C-means and region growing based classification of tumor from mammograms using hybrid texture feature. *J Comput Sci* 29:34–45. <https://doi.org/10.1016/j.jocs.2018.09.015>
- Saeedi S, Rezayi S, Keshavarz H, Niakan-Kalhari SR (2023) MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Med Inform Decis Mak* 23(1):1–17. <https://doi.org/10.1186/s12911-023-02114-6>
- Safdar M, Kobaisi S, Zahra F (2020) A comparative analysis of data augmentation approaches for magnetic resonance imaging (MRI) scan images of brain tumor. *Acta Informatica Medica* 28(1):29. <https://doi.org/10.5455/aim.2020.28.29-36>
- Salem-Ghahfarrokhi S, Khodadadi H (2020) Human brain tumor diagnosis using the combination of the complexity measures and texture features through magnetic resonance image. *Biomed Signal Process Control* 61:102025. <https://doi.org/10.1016/j.bspc.2020.102025>
- Sánchez-Cauce R, Pérez-Martín J, Luque M (2021) Multi-input convolutional neural network for breast cancer detection using thermal images and clinical data. *Comput Methods Programs Biomed*. <https://doi.org/10.1016/j.cmpb.2021.106045>
- Sannasi-Chakravarthy SR, Rajaguru H (2022) Automatic detection and classification of mammograms using improved extreme learning

- machine with deep learning. *IRBM* 43(1):49–61. <https://doi.org/10.1016/j.irbm.2020.12.004>
- Saravana-Kumar NM, Hariprasath K, Tamilselvi S, Kavinya A, Kaviyavarshini N (2021) Detection of stages of melanoma using deep learning. *Multimed Tools Appl* 80(12):18677–18692. <https://doi.org/10.1007/s11042-021-10572-1>
- Saric M, Russo M, Stella M, Sikora M (2019) CNN-based method for lung cancer detection in whole slide histopathology images. In: 2019 4th International Conference on Smart and Sustainable Technologies, SpliTech 2019, pp. 14–17. <https://doi.org/10.23919/SpliTech.2019.8783041>
- Scarpace L et al. (2016) The cancer genome atlas glioblastoma multi-forme collection (TCGA-GBM) (version 4). The Cancer Imaging Archive
- Selvapandian A, Nagendra-Prabhu S, Sivakumar P, Jagannadha-Rao DB (2022) Lung cancer detection and severity level classification using sine cosine sail fish optimization based generative adversarial network with CT images. *Comput J* 65(6):1611–1630. <https://doi.org/10.1093/comjnl/bxab141>
- Senan EM, Jadhav ME (2021) Analysis of dermoscopy images by using ABCD rule for early detection of skin cancer. *Glob Trans Proc* 2(1):1–7. <https://doi.org/10.1016/j.glt.2021.01.001>
- Shafi ASM, Rahman MB, Anwar T, Halder RS, Kays HME (2021) Classification of brain tumors and auto-immune disease using ensemble learning. *Inform Med Unlocked* 24:100608. <https://doi.org/10.1016/j.imu.2021.100608>
- Shahsavari A, Khatibi T, Ranjbari S (2023) Skin lesion detection using an ensemble of deep models: SLDED. *Multimed Tools Appl* 82(7):10575–10594. <https://doi.org/10.1007/s11042-022-13666-6>
- Shahzadi I, Tang TB, Meriadeau F, Quyyum A (2018) CNN-LSTM: cascaded framework for brain tumour classification. In: 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), IEEE, pp. 633–637. <https://doi.org/10.1109/IECBES.2018.8626704>
- Sharma S, Mehra R (2020) Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight. *J Digit Imaging* 33(3):632–654. <https://doi.org/10.1007/s10278-019-00307-y>
- Sheeba A, Santhosh-Kumar P, Ramamoorthy M, Sasikala S (2023) Microscopic image analysis in breast cancer detection using ensemble deep learning architectures integrated with web of things. *Biomed Signal Process Control* 79(P2):104048. <https://doi.org/10.1016/j.bspc.2022.104048>
- Shetty B, Fernandes R, Rodrigues AP, Vijaya P (2022) Brain tumor detection using machine learning and convolutional neural network. In: 2022 International Conference on Artificial Intelligence and Data Engineering (AIDE), IEEE, pp. 86–91. doi: <https://doi.org/10.1109/AIDE57180.2022.10060254>
- Shim SO, Alkinani MH, Hussain L, Aziz W (2022) Feature ranking importance from multimodal radiomic texture features using machine learning paradigm: a biomarker to predict the lung cancer. *Big Data Res* 29:100331. <https://doi.org/10.1016/j.bdr.2022.100331>
- Shimanto SA, Hosain MK, Biswas SP, Islam MS (2023) Brain tumor detection and classification by SVM algorithm and performance analysis through CNN approach. In: 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE), IEEE, pp. 1–6. <https://doi.org/10.1109/ECCE57851.2023.10101618>
- Suckling J, Parker J, Dance D, Astley S, Hutt I (2015) Mammographic Image Analysis Society (MIAS) database v1.21. Apollo - University of Cambridge Repository., 2015. <https://www.repository.cam.ac.uk/handle/1810/250394> accessed 21 Jul 2023
- Talo M, Yildirim O, Baloglu UB, Aydin G, Acharya UR (2019) Convolutional neural networks for multi-class brain disease detection using MRI images. *Comput Med Imaging Graph* 78:101673. <https://doi.org/10.1016/j.compmedimag.2019.101673>
- Tharwat A (2018) Classification assessment methods. *Appl Comput Inform.* <https://doi.org/10.1016/j.aci.2018.08.003>
- The Indian Express (2019) World brain tumour day 2019: know the symptoms, risk factors and treatment. <https://indianexpress.com/article/lifestyle/health/world-brain-tumour-day-2019-symptoms-risk-factors-treatment-5770587> accessed 7 Jun 2020
- Toğaçar M, Ergen B, Cömert Z (2020) BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model. *Med Hypotheses* 134:109531. <https://doi.org/10.1016/j.mehy.2019.109531>
- Tschandl P, Rosendahl C, Kittler H (2018a) The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 5(1):180161. <https://doi.org/10.1038/sdata.2018.161>
- Tschandl P, Rosendahl C, Kittler H (2018b) Data descriptor: the HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 5:1–9. <https://doi.org/10.1038/sdata.2018.161>
- Tumpa PP, Kabir MA (2021) An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features. *Sens Int* 2:100128. <https://doi.org/10.1016/j.sintl.2021.100128>
- United Nations Conference on Trade and Development (2022) UN list of least developed countries. UNCTAD. <https://unctad.org/topic/least-developed-countries/list>
- Ur-Rehman K, Li J, Pei Y, Yasin A, Ali S, Mahmood T (2021) Computer vision-based microcalcification detection in digital mammograms using fully connected depthwise separable convolutional neural network. *Sensors*. <https://doi.org/10.3390/s21144854>
- Vaiyapuri T, Liyakathunisa, Alaskar H, Parvathi R, Pattabiraman V, Hussain A (2022) Cat swarm optimization-based computer-aided diagnosis model for lung cancer classification in computed tomography images. *Appl Sci (switzerland)*. <https://doi.org/10.3390/app12115491>
- Vaka AR, Soni B, Reddy SK (2020) Breast cancer detection by leveraging machine learning. *ICT Express* 6(4):320–324. <https://doi.org/10.1016/j.icte.2020.04.009>
- Valvano G et al (2019) Convolutional neural networks for the segmentation of microcalcification in mammography imaging. *J Healthc Eng.* <https://doi.org/10.1155/2019/9360941>
- Vankdothu R, Hameed MA (2022) Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning. *Meas Sens* 24:100440. <https://doi.org/10.1016/j.measen.2022.100440>
- Vijayarajeswari R, Parthasarathy P, Vivekanandan S, Basha AA (2019) Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement (lond)* 146:800–805. <https://doi.org/10.1016/j.measurement.2019.05.083>
- Vineeth J, Hemanth S, Rao CV, Pavankumar N, Javanna HS, Janardhan CN (2022) Skin cancer detection using deep learning. In: 2022 4th International Conference on Cognitive Computing and Information Processing, CCIP 2022, no. Icears, pp. 1724–1730. <https://doi.org/10.1109/CCIP57447.2022.10058685>
- Virupakshappa, Amarapur B (2020) Computer-aided diagnosis applied to MRI images of brain tumor using cognition based modified level set and optimized ANN classifier. *Multimed Tools Appl* 79(5–6):3571–3599. <https://doi.org/10.1007/s11042-018-6176-1>
- Wahba MA, Ashour AS, Guo Y, Napoleon SA, Abd MM (2018) Computer methods and programs in biomedicine a novel cumulative level difference mean based GLDM and modified ABCD features ranked using eigenvector centrality approach for four skin

- lesion types classification. *Comput Methods Programs Biomed* 165:163–174. <https://doi.org/10.1016/j.cmpb.2018.08.009>
- Wang Z, Xin J, Sun P, Lin Z, Yao Y, Gao X (2018) Improved lung nodule diagnosis accuracy using lung CT images with uncertain class. *Comput Methods Programs Biomed* 162:197–209. <https://doi.org/10.1016/j.cmpb.2018.05.028>
- World Health Organization (2019) Global cancer observatory. Malaysia Cancer Statistics. <https://gco.iarc.fr/> accessed 19 May 2023
- World Health Organization International Agency for Research on Cancer (2020) The Global Cancer Observatory—all cancers. International Agency for Research on Cancer - WHO, vol. 419, pp. 199–200
- Woźniak M, Połap D, Capizzi G, Lo-Sciuto G, Kośmider L, Frankiewicz K (2018) Small lung nodules detection based on local variance analysis and probabilistic neural network. *Comput Methods Programs Biomed* 161:173–180. <https://doi.org/10.1016/j.cmpb.2018.04.025>
- Yan F, Huang H, Pedrycz W, Hirota K (2023) Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms. *Expert Syst Appl* 227:120282. <https://doi.org/10.1016/j.eswa.2023.120282>
- Yu K-H et al (2020) Classifying non-small cell lung cancer types and transcriptomic subtypes using convolutional neural networks. *J Am Med Inform Assoc* 27(5):757–769. <https://doi.org/10.1093/jamia/ocz230>
- Zakareya S, Izadkhah H, Karimpour J (2023) A new deep-learning-based model for breast cancer diagnosis from medical images. *Diagnostics* 13(11):1944. <https://doi.org/10.3390/diagnostics13111944>
- Zeng W, Liao Y, Chen Y, Ying-Diao Q, Ying-Fu Z, Yao F (2023) Research on classification and recognition of the skin tumors by laser ultrasound using support vector machine based on particle swarm optimization. *Opt Laser Technol* 158:108810. <https://doi.org/10.1016/j.optlastec.2022.108810>
- Zhang N, Cai YX, Wang YY, Tian YT, Wang XL, Badami B (2020) Skin cancer diagnosis based on optimized convolutional neural network. *Artif Intell Med* 102:101756. <https://doi.org/10.1016/j.artmed.2019.101756>
- Zhao J, Chen T, Cai B (2022) A computer-aided diagnostic system for mammograms based on YOLOv3. *Multimed Tools Appl* 81(14):19257–19281. <https://doi.org/10.1007/s11042-021-10505-y>
- Zhou H et al (2018) Diagnosis of distant metastasis of lung cancer: based on clinical and radiomic features. *Transl Oncol* 11(1):31–36. <https://doi.org/10.1016/j.tranon.2017.10.010>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.