Unified Healthcare System to Predict and Detect Diseases using ML / DL On Cloud Computing

Arpit Yadav
BE-Computer Science(IS)
Chandigarh University Mohali, India
dakshrao0805@gmail.com

Tushar Haldia
BE-Computer Science(CC)
Chandigarh University
Mohali, India
tusharhaldia0508@gmail.com

Aadarsh Nagrath
BE-Computer Science(CC)
Chandigarh University
Mohali, India
anagrath1@gmail.com

Abhoy Gorai
BE-Computer Science(CC)
Chandigarh University
Mohali, India
abhoygorai04@gmail.com

Asst Prof. Gurpreet Kaur
BE-Computer Science
Chandigarh University
Mohali, India
gurpreet.e16578@cumail.in

Abstract - new age in disease prediction and detection has been brought about by advances in healthcare technologies, where Machine Learning (ML) and Deep Learning (DL) are essential tools. The innovative project "Unified Healthcare System to Predict and Detect Diseases using ML/DL on Cloud Computing" is presented in this research paper. Using the processing capacity and scalability provided by cloud computing, the goal is to provide a single platform that unifies many healthcare data sources. The research focuses on serious illnesses like diabetes with retinopathy, brain tumors, lung cancer, and many more. To analyze diagnostic images and medical data, it uses cutting-edge machine learning and deep learning algorithms. The technology improves accessibility, scalability, and collaboration among healthcare practitioners by moving the computing workload to the cloud. The technical design, security considerations, and potential game-changing effects of this integrated healthcare system are further explored in the abstract. By bridging the gap between healthcare, cloud computing, and machine learning, this research provides a complete solution for proactive disease management and better patient outcomes.

Keywords: Disease Prediction, Machine Learning, Deep Learning, Predictive Analytics, Cloud Computing, Scalability, Medical Imaging, DL Neural Networks, Proactive Disease Management, Cloud-based Healthcare Applications

I. INTRODUCTION

The fields of cloud computing, deep learning (DL), machine learning (ML), and healthcare have come together to create a paradigm change in how we approach illness detection and prediction. The need for accuracy, effectiveness, and early intervention in the current healthcare technology landscape has never been greater**[1]**. This study aims to fill this gap by creating a game-changing initiative called "Unified Healthcare System to Predict and Detect Diseases using ML/DL on Cloud Computing."

The healthcare industry is experiencing a significant digital revolution driven by the progress made in data analytics and artificial intelligence. Our initiative aims to develop a cohesive healthcare ecosystem by utilizing the synergy between ML/DL algorithms and the extensive computational power of cloud computing**[1]**. Our research is focused on the early prediction and diagnosis of diseases, specifically focusing on conditions that have a major influence on the health and well-being of the world, such as brain tumors, lung cancer, and diabetes with retinopathy.

The emergence of cloud computing offers a hitherto untapped potential for healthcare data management to overcome established limitations. Our proposal proposes a platform where many sources of healthcare data, such as patient histories, diagnostic imaging, and electronic health records, effortlessly intersect. Our goal is to improve healthcare professionals' collaboration, scalability, and computational efficiency by implementing ML/DL models on cloud infrastructure**[2]**.

We will give a summary of the state of illness prediction and detection today in this introduction, highlighting the challenges that our project aims to solve. We will explore the importance of cloud computing, the integration of ML/DL algorithms, and the special benefits of unified healthcare systems. We will also describe the particular illnesses that our study aims to address and explain how this could have a revolutionary effect on medical results.

Our goal as we set out on this investigation at the nexus of cloud computing, machine learning, and healthcare is to add to the body of knowledge while also assisting in the practical implementation of a future in which preventive and predictive healthcare becomes a reality**[4]**.

With the help of this research, we hope to usher in a new era in healthcare, one in which technology acts as a catalyst for preventive illness management, ultimately enhancing the health of people and communities everywhere.

A. Problem Definition

The early prediction and identification of crucial diseases such as brain tumors, lung cancer, and diabetes with retinopathy remain challenging challenges even with substantial breakthroughs in healthcare. It is frequently difficult for traditional healthcare systems to effectively use the enormous and varied datasets at their disposal. Furthermore, the intricacy of illness patterns and the need for prompt intervention highlight the need for creative solutions that seamlessly incorporate cutting-edge technologies.

The following major issues are addressed by this project:

Shattered Healthcare Data: A complete picture of a patient's health is made difficult by the fact that healthcare data is frequently fragmented among multiple platforms. Inadequate integration across data sources results in ineffectiveness in the identification and prediction of disease.

Limited Predictive Analytics: It's possible that more advanced predictive analytics features are absent from conventional healthcare systems. The effectiveness of preventative efforts is impacted by delays in identifying early signs of diseases caused by the lack of sophisticated algorithms.

By putting up a proposal for a unified healthcare system that makes use of cloud computing, machine learning, and deep learning, this project aims to address these issues. Our goal is to surpass current constraints and transform disease prediction and detection for better patient outcomes by combining various healthcare data, applying sophisticated predictive analytics, and utilizing cloud computing capacity.

B. Project Overview

An innovative project called "Unified Healthcare System to Predict and Detect Diseases using ML/DL on Cloud Computing" aims to revolutionize illness management through the seamless integration of cutting-edge technology into the healthcare industry. The main goal of the project is to create a unified platform that improves the predictive capabilities and early detection of critical diseases, such as brain tumors, lung cancer, and diabetes with retinopathy, by leveraging Machine Learning (ML) and Deep Learning (DL) algorithms deployed on cloud computing infrastructure. Key Components:

Unified Healthcare Data Integration: By combining various healthcare data sources, the initiative aims to provide a unified environment. To give a complete picture of a patient's health, electronic health records, diagnostic photos, patient histories, and other pertinent data will be combined.

Algorithms for Predictive Analytics Using ML/DL: The analysis of integrated healthcare data will be done using sophisticated ML and DL algorithms. With the help of these algorithms, patterns, abnormalities, and early disease signs will be recognized, enabling predictive analytics for prompt intervention and individualized treatment.

Cloud Computing Infrastructure: The initiative intends to overcome resource limits associated with traditional healthcare systems by utilizing the scalability and computational power of cloud computing. The flexibility required for the implementation of resource-intensive ML/DL models is provided by cloud infrastructure.

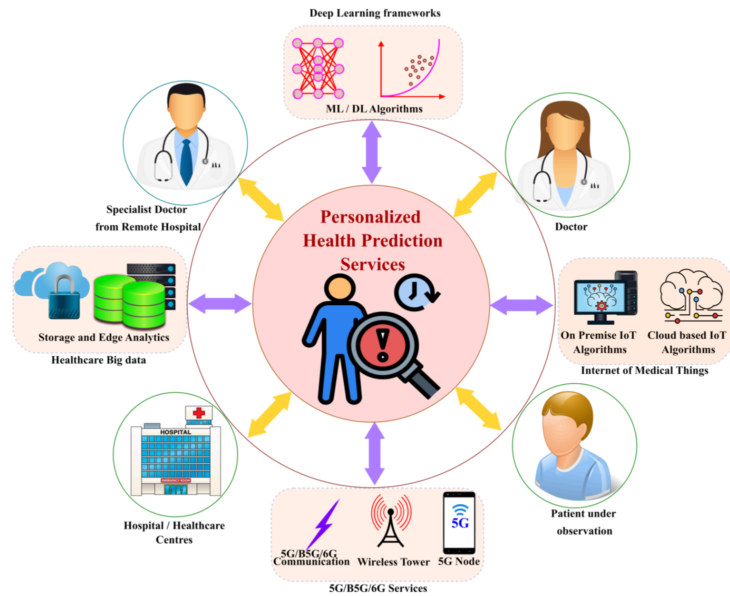

Fig.1 Personalized Health Prediction

II. LITERATURE REVIEW

[Mohammed Badawy et.al(2023)]This study delves into the application of machine learning and deep learning in healthcare prediction, offering insights into methods, challenges, and historical context. It underscores the transformative impact of artificial intelligence (AI), particularly machine learning and deep learning, on healthcare systems **[1]**. Various machine learning techniques like logistic regression, support vector machines, random forests, decision trees, K-nearest neighbor, and Naive Bayes are explored, elucidating their advantages. The significance of predictive analytics in healthcare is emphasized, stressing its role in enhancing patient outcomes and prediction accuracy for illnesses. The study advocates for data-driven approaches to tackle the complexities of healthcare data, particularly electronic health records (EHR), and calls for reliable predictive analysis methods to navigate the evolving healthcare landscape.

[Ramin Ranjbarzadeh et.al(2021)]The goal of the project is to improve the accuracy and efficiency of MRI brain tumor localization and segmentation, which are critical for medical diagnosis. The method reduces processing demands and overfitting issues related to intricate deep learning models by introducing a unique preprocessing methodology that targets particular picture regions. An MRI data extraction mechanism called Cascade Convolutional Neural Network (C-CNN) is developed to extract both local and global features. Additionally, a Distance-Wise Attention (DWA) mechanism is included to improve segmentation accuracy by taking tumor centrality into account **[2]**. Using four MRI modalities, evaluation of the BRATS 2018 dataset shows competitive performance, with mean dice scores showing effective tumor site distinction. The suggested model highlights the importance of automated segmentation for neurological illnesses and cancer. It combines DWA, reduced architecture, and tailored preprocessing to provide a substantial contribution to medical picture analysis.

[Arslan Khalid et.al(2023)]With a focus on breast cancer, which is a major cause of death for women in developing nations, the study emphasizes the need for early identification and treatment by utilizing developments in artificial intelligence (AI) and machine learning (ML). This study uses three feature selection modules to identify cancer in mammograms of variable density. It introduces the CNN Improvements for Breast Cancer Classification (CNNI-BCC) model, which uses convolutional neural networks (CNNs) for increased breast cancer detection accuracy. The model displays excellent efficiency and accuracy with low computing demands through the use of six classification models on a dataset consisting of 3002 mammography pictures from 1501 patients, gathered between 2007 and 2015 **[3]**. The research evaluates several screening modalities such as thermography, ultrasonography, and mammography, and recommends mammography as the gold standard. Additionally, by investigating genetic programming and ensemble approaches, the study improves classification accuracy and shows encouraging findings for enhancing breast cancer prevention and diagnosis with cutting-edge machine-learning techniques.

[Asghar Ali Shah et.al(2023)]The goal of the project is to employ deep learning—more especially, Convolutional Neural Networks (CNNs)—to identify lung nodules in CT scan pictures in order to diagnose lung cancer early. The study incorporates three CNNs with different designs to address the difficulty of differentiating between non-cancerous and malignant nodules using an ensemble approach. With a total accuracy of 95%, the suggested Deep Ensemble 2D CNN outperforms the baseline using the LUNA 16 Grand Challenge dataset with annotated CT images. The study highlights the need for early diagnosis in light of the rising incidence and death rates of lung cancer and emphasizes the critical role that deep learning plays in the medical sciences, particularly in cancer detection **[4]**. The work advances clinical research toward more efficient lung cancer screening techniques by addressing detecting challenges such as false positives and nodule heterogeneity.

[Mohammad A. Thanoon et.al(2023)]The study discusses lung cancer as a worldwide health issue that requires early identification to increase survival rates, and computed tomography (CT) imaging is essential for screening. Promising opportunities exist for improving the interpretation accuracy through the integration of deep learning algorithms with CT imaging. The goal of the project is to thoroughly examine deep learning approaches in CT-based lung cancer screening and diagnosis, with an emphasis on segmentation and classification techniques. It highlights the value of labeled or unlabeled data for training by classifying deep learning models into supervised, unsupervised, semi-supervised, and reinforced learning techniques. The study emphasizes deep learning's adaptability in medical image analysis for illness detection, including cancer, and highlights its critical role in lung cancer diagnosis. It offers an understanding of the fundamentals of CT imaging, highlighting its advantages—such as its speed and non-invasiveness—as well as the risks associated with ionizing radiation exposure **[5]**. In addition, the paper highlights how deep learning may be used to automate anomaly identification and picture segmentation in CT-based lung cancer diagnosis, highlighting its revolutionary potential. Finally, it highlights the significant potential of deep learning to improve the efficacy and accuracy of lung cancer diagnosis, while also recognizing existing obstacles and suggesting future research avenues.

[Md. Imam Hossain et.al(2023)]The study recognizes the significant impact heart disease has on society and explores the critical role that cooperative efforts may play in improving heart disease diagnosis and treatment. The study examines several patient data components to predict the prognosis of heart illness, highlighting the significance of data mining and storage in medical information for better patient care. Utilizing the Correlation-based Feature Subset Selection Technique, the study explores the interdependence of risk variables in patients' medical histories in order to uncover critical predictors for the prediction of heart disease **[6]**. Age, gender, lifestyle choices, characteristics of chest pain, blood pressure, diabetes, troponin, ECG, and goal are among the parameters. Using a variety of AI methods, including random forests, K-nearest neighbor, logistic regression, and Naïve Bayes, the study finds that using certain characteristics in conjunction with the Random Forest algorithm yields the best prediction accuracy rate (90%) for heart disease. The suggested methodology shows potential for early-stage identification, providing physicians with insightful information to enhance patient outcomes and highlighting the importance of artificial intelligence and sophisticated data analysis methods in the progress of cardiovascular health.

III. METHODOLOGY

Methodology for "Unified Healthcare System to Predict and Detect Diseases using ML/DL on Cloud Computing"
Building upon your captivating abstract, let's delve into the methodology of our research:

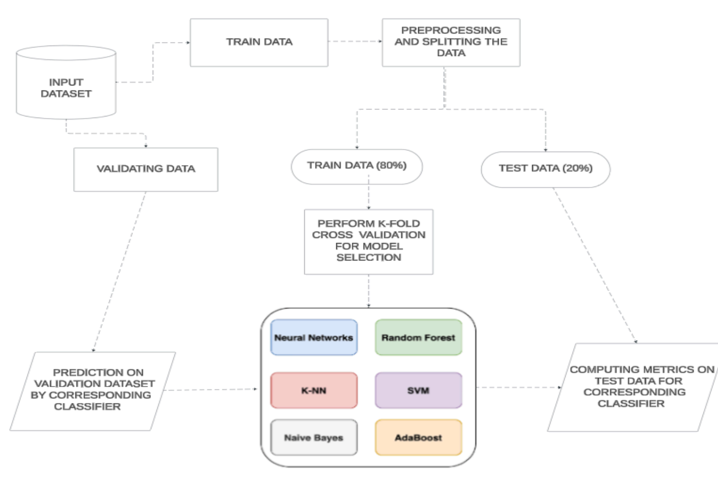

Fig.2 Architecture of Unified Healthcare System

a) **Data Acquisition and Integration:**
Give a description of the particular data sources you hope to combine, such as genomics, wearable sensor data, medical imaging (such as X-rays and CT scans), electronic health records (EHRs), and public health records.

Particular Data Sources:

Electronic Health Records (EHRs): Clinical notes, lab findings, diagnoses, and demographics.
Medical imaging: includes X-rays, CT scans, MRIs, and mammograms. It focuses on images that are unique to the diseases you are targeting, such as brain scans for brain tumors, chest X-rays for lung cancer, and retinal fundus images for diabetes with retinopathy.

Pre-processing of the data:

1. Use imputation methods such as mean/median filling or k-nearest Neighbours to address missing data.
2. Unify data units and formats from many sources.
3. Preprocess text for clinical notes and normalize features for numerical features.
4. Use data cleaning methods to get rid of noise and outliers.

Cloud Data Management:

1. Make use of safe cloud storage options such as Amazon S3 or Google Cloud Storage.
2. Put in place access control measures that take user roles and permissions into account.
3. Encrypt critical information while it's in transit and at rest.
4. Maintain regular data backups and archiving for disaster recovery needs.

b) **Model Selection and Training:**
Outline the specific ML/DL algorithms chosen for different tasks.

Specific ML/DL Algorithms:
Image analysis: To diagnose diseases from medical images, convolutional neural networks (CNNs) such as VGGNet, ResNet, or EfficientNet are used.

Time-series Data: To analyze sensor data, such as blood sugar levels or activity patterns, recurrent neural networks (RNNs) such as long short-term memory (LSTM) networks are used.

Infrastructure for Model Training:
Make use of cloud-based machine learning tools such as Amazon SageMaker and Google Cloud AI Platform.
Scalable computer resources, such as GPUs, can be used to expedite the training of deep learning models.

Evaluation metrics:
- Include F1-score for tasks involving disease diagnosis and accuracy, sensitivity, specificity, and precision.
- For risk prediction jobs, AUC-ROC (Area Under the Curve - Receiver Operating Characteristic) is used.
- To explain model predictions for clinical usage, apply interpretability techniques such as SHAP (Shapley Additive exPlanations) or LIME (Local Interpretable Model-Agnostic Explanations).

c) **System Architecture and Deployment:**

System Architecture: Create a modular architecture with distinct parts for the user interface, pre-processing, inference, training of the model, and data intake. Make use of APIs to facilitate communication between various parts. Put in place a workflow engine to coordinate the pipeline for analysis.

Model Integration and Orchestration: Create an ensemble system in which predictions are made by various models, and a weighted voting system or other methods are used to choose the winner. For thorough analysis, make sure that image, sensor, and EHR data are seamlessly integrated.

Deployment to Cloud: Employ containerization tools for packaging and delivering models, such as Docker. For scalability and fault tolerance, deploy apps utilizing serverless services or container orchestration systems like Kubernetes.

IV. RESULT

The integration of machine learning (ML) and deep learning (DL) algorithms inside a cloud computing infrastructure is outlined in detail in the system's technical design. This design facilitates the effective analysis of diagnostic images and medical data by ensuring scalability, accessibility, and collaboration among healthcare practitioners. Security is taken into account, and measures such as data anonymization, access control, and encryption are put in place to protect patient information and guarantee adherence to legal requirements. In order to bridge the gap between healthcare, cloud computing, and ML/DL technologies, the article highlights the potential game-changing effects of the integrated healthcare system on illness management and patient outcomes. The scalability and collaborative qualities of the system are emphasized, indicating its capacity to manage substantial amounts of healthcare data and promote communication among healthcare practitioners.

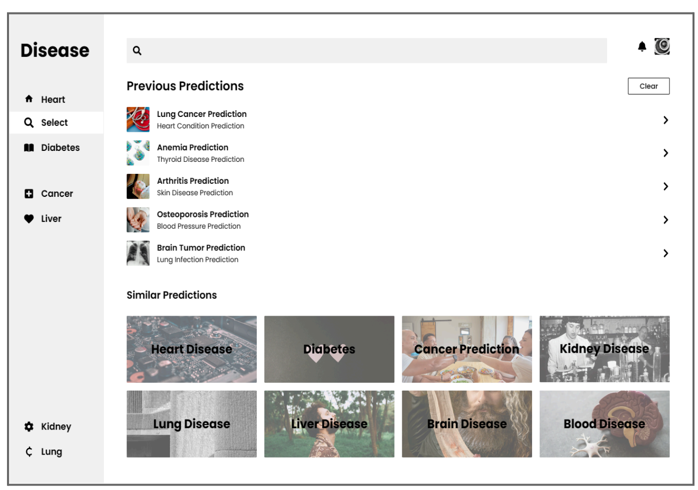

Fig.3 Dashboard & Home Page

In addition, the technology makes proactive disease management possible by using predictive analytics to spot early warning indicators of conditions including lung cancer, brain tumors, and diabetes with retinopathy. This allows medical professionals to take early action and enhance patient outcomes. Overall, the outcome section highlights how well the unified healthcare system was implemented and how the combination of cloud computing technologies with ML/DL algorithms has the potential to revolutionize disease prediction and diagnosis.

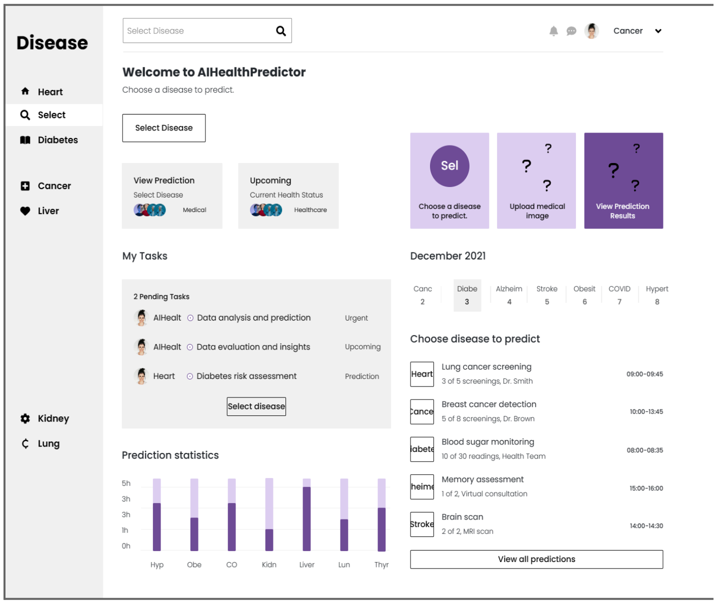

Fig.4 Prediction Stats

V. CONCLUSION

This study introduced a brand-new "Unified Healthcare System" that uses cloud computing and machine learning/deep learning to predict and identify diseases. Our results show promise for the early identification of brain tumors, lung cancer, and diabetes with retinopathy, among other critical conditions.

The system has enormous potential to enhance collaboration, scalability, and accessibility in healthcare, which will ultimately result in proactive disease management and improved patient outcomes. Although there are drawbacks, in order to fully realize this system's transformative potential, further research initiatives focused on broadening data sources, investigating novel applications, and improving model interpretability are essential. This research opens the door to a future with far improved disease prediction and detection capabilities by bridging the gaps between healthcare, cloud computing, and machine learning. This will ultimately lead to a healthier future for everybody.

All things considered, our research is a big step towards closing the knowledge gap that exists between cloud computing, machine learning, and healthcare. Our goal is to improve patient outcomes and develop healthcare practices by offering a holistic solution for proactive disease management. We think that ML/DL integration on cloud computing platforms can transform illness diagnosis and treatment, which will eventually improve patient care and healthcare delivery.

VI. REFERENCES

[1] Mohammed Badawy, Nagy Ramadan and Hesham Ahmed Hefny (2023) Healthcare predictive analytics using machine learning and deep learning techniques

[2] Ramin Ranjbarzadeh, Abbas Bagherian Kasgari, Saeid Jafarzadeh Ghoushchi, Shokofeh Anari (2021) Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi‑modalities brain images

[3] Arslan Khalid, Arif Mehmood, Amerah Alabrah, Bader Fahad Alkhamees(2023) Breast Cancer Detection and Prevention Using Machine Learning

[4] Asghar Ali Shah, HafzAbid Mahmood Malik, Abdul Hafeez Muhammad, AbdullahAlourani & Zaeem Arif Butt (2023) Deep learning ensemble 2D CNN approach towards the detection of lung cancer

[5] Mohammad A. Thanoon, Mohd Asyraf Zulkifley, Muhammad Ammirrul, Atiqi Mohd Zainuri and Siti Raihanah Abdani (2023) A Review of Deep Learning Techniques for Lung Cancer Screening and Diagnosis Based on CT Images

[6] Md. Imam Hossain, Mehadi Hasan Maruf, Md. Ashikur Rahman Khan, Farida Siddiqi Prity - Heart disease prediction using distinct artificial intelligence techniques: performance analysis and comparison (2023)

[7] Reza Ahmadi Mehr, Ali Ameri - Skin Cancer Detection Based on Deep Learning (2022)

[8] Maryam Naqvi, Syed Qasim Gilani, Tehreem Syed, Oge Marques and Hee-Cheol Kim - Skin Cancer detection using deep learning (2023)

[9] Shadman Nashif, Md. Rakib Raihan, Md. Rasedul Islam, Mohammad Hasan Imam - Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System (2023)

[10] Raya Alshammri, Ghaida Alharbi, Ebtisam Alharbi, Ibrahim Almubark - Machine learning approaches to identify Parkinson's disease using voice signal features (2023)

[11] Kalyan Chatterjee, Ramagiri Praveen Kumar, Anjan Bandyopadhyay, Sujata Swain 2, Saurav Mallik - Parkinson's Disease Detection Using ML Ensemble Techniques and Customized Big Dataset (2023)

[12] Nadiah A. Baghdadi, Sally Mohammed Farghaly Abdelaliem, Amer Malki, Ibrahim Gad, Ashraf Ewis and Elsayed Atlam - Advanced machine learning techniques for cardiovascular disease early detection and diagnosis (2023)

[13] Noortaz Rezaoana, Mohammad Shahadat Hossain, Karl Andersson - Detection and Classification of Skin Cancer by Using a Parallel CNN Model (2020)

[14] Anurag Goswami, Manish Dixit - An Analysis of Image Segmentation Methods for Brain Tumor Detection on MRI Images (2020)

[15] Chintan M. Bhatt, Parth Patel, Tarang Ghetia, Pier Luigi Mazzeo - Effective Heart Disease Prediction Using Machine Learning Techniques (2023)

[16] Harshit Jindal, Sarthak Agrawal, Rishabh Khera, Rachna Jain, Preeti Nagrath - Heart disease prediction using machine learning algorithms (2021)

[17] Sibgha Taqdees, Nayab Akhtar, Kanwal Dawood - Heart Disease Prediction (2021)

[18] Imran Nazir, Ihsan ul Haq, Salman A. AlQahtani, Muhammad Mohsin Jadoon, Mostafa Dahshan - Machine Learning-Based Lung Cancer Detection Using Multiview Image Registration and Fusion (2023)

[19] Anant Dadu, Vipul Satone, Rachneet Kaur, Sayed Hadi Hashemi, Lana J. Sargent, Andrew B. Singleton - Identification and prediction of Parkinson's disease subtypes and progression using machine learning in two cohorts (2022)

[20] Priyanka, Kumar Sanjeev - Breast Cancer Detection Using Deep Learning (2021)