

MedXPert: A Novel Approach of Integrating G-AI With Medical Field by Combining Models and Making It Easily Accessible for Public Within their Finger-Tips

Sai Tharun Peram¹, L Sree Nidhi², Hashmmath Shaik³, Taathvika M⁴

saitharun_peram@srmap.edu.in, sreenidhi_l@srmap.edu.in,
hashmmath_shaik@srmap.edu.in, taathvika_m@srmap.edu.in.

Under the Guidance of Dr. Rajiv Senapati⁵ (rajiv.s@srmap.edu.in)

SRM University AP, Amaravati, Andhra Pradesh, 522240, India.

Abstract

Generative AI has now become a booming software in the present digital and modernized world. This has proven its worth by showcasing excellent and reliable results in multiple fields, the one field that needs concentration is the Medical and Healthcare Industry as it is very complex and people rely on it. Along with G-AI there are so many advancements in Machine Learning and Deep Learning Architectures to accurately predict and classify the medical reports for public to evaluate and practitioners to analyse making their job simple. The model that we are discussing in the report is the advancement of the existing models with the integration of Generative-AI and incorporating data pre-processing and models like Image Enhancement, Image Classification, and combining it with AI, Machine Learning, Deep Learning, CNN, NLP, XAI (Explainable AI), Visual Transformers, GANs, LLMs, Web Technologies, and API References to deliver tailored diet, exercise, remedies, medication, and consultation advice based on comprehensive medical history, and Customer Data. The model will enhance and analyse MRI scans, CT scans, and other medical reports like EHR to extract valuable insights like physiological conditions, anatomical characteristics, risk of tumours, and disease risk factors and generate informative outputs for users using the above models and technologies, then process it to G-AI to tailor reports and answer queries based on users' request. With the enhancement of scans and reports we were able to achieve higher accuracy by evaluating using Precision, Recall, F1-Score. Which is compared with the state-of-the-art existing models.

Keywords: Generative-AI, Medical, Healthcare, Reports, Data Pre-processing, Image Enhancement, Image Classification, AI, Machine Learning, Deep Learning, CNN, NLP, Visual Transformers, GANs, LLMs, Web Technologies, API, Diet, Medication, Consultation, MRI, CT, Tumours, Disease, EHR.

1. Introduction

With the advancement in the Generative-AI (GEN-AI) field, many models were brought to light such as ChatGPT (Personalized GEN-AI Content Generator), DALL-E (Personalized GEN-AI Image Generator from text), Google Gemini (Google's ChatGPT version) and many more. These models encompass Audio Generation, 3D-Image and Video Generator, and Code Generation like Blackbox.ai, with more to come. The numerous platforms built have the idea of pushing GEN-AI to its greatest potential. Now, these advancements are utilized in various fields such as Medical, Financial, Gaming, Software Industries, Urban Planning, Content generation for social media, and other sectors.

These GEN-AI models operate through the integration of complex techniques like Generative Adversarial Networks (GANs), Transformers, and Variational Autoencoders. The methodology for GEN-AI Chatbot (for temporary use) is discussed below, along with a detailed explanation of GEN-AI, Large Language Models (LLMs), and their functionality, including various examples related to existing market models.

This project aims to build a personalized GEN-AI model focused on the health and medical industry, generating ideas, remedies, diet plans, medication, prevention, and more. The report will delve further into these topics, with the results being entirely user-based, tailored to individual aspects derived from their medical records. The next segment involves analyzing medical reports and scans like MRI Scans, CT scans, User Medical Reports, and Users' info using Deep Learning models, Machine Learning models, Natural Language Processing (NLP), Convolutional Neural Networks (CNN), Artificial Intelligence (AI), and other technologies. Along with those technologies and methods Data Pre-Processing and Image Enhancement techniques have also made significant contribution in medical field.

The Medical and Healthcare industry has undergone a revolution with the advent of AI and Machine Learning (ML), widely employed in tumor detection in areas such as Lungs, Brain, Breast, Liver, and Intestinal. Techniques like NLP, CNN, Deep Learning, and Visual Transformers are utilized in various medical fields for scanning and analyzing medical reports, diabetes detection using retinal images, extracting data from images, and generative aging of Brain Images within a given timespan for the same user.

These advancements have indeed revolutionized the Healthcare and Medical Industry, simplifying tasks that once required professional assistance. With ongoing research and developments in tumor detection and report scanning, the field must explore new technologies to integrate and scale up for a better, efficient, reliable, and user-friendly model. Among the available technologies, our team has

chosen GEN-AI. Under our mentor's guidance, we developed and implemented this idea.

The model discussed in the report comprises various technologies and features. The main interface is a chat interface with options like writing text, importing media, and audio-transcription aimed at engaging user interaction. Users can upload medical reports to be evaluated using the mentioned techniques for medical image analysis including data pre-processing, image enhancement techniques. Then the output data is generated in a detailed manner in text form with evaluating prompts and taking the efficient prompt. Then GEN-AI generates content based on user inquiries for deeper insights and more precise recommendations like mentioned before utilizing the techniques mentioned earlier. Further details on specifications, usage, models, technologies, accuracy, and distinctions from other models are elaborated in subsequent sections.

2. Literature review

The field of generative AI focuses on developing algorithms and models capable of generating synthetic data that closely resemble real-world data. The ability to generate realistic and novel data has immense implications across multiple industries, including entertainment, healthcare, finance, etc. Generative AI has opened new avenues for applications such as image synthesis, text generation, music composition, and even human-like chatbots [1]. Generative artificial intelligence (AI) has displayed significant potential in advancing the medical field. Medical imaging plays a crucial role in modern healthcare, aiding in early diagnosis and treatment planning. Various imaging modalities like X-ray, computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound are routinely used to diagnose and monitor a wide array of medical conditions. These advancements have led to more precise diagnoses, treatment plans, and outcome assessments, enhancing medical practice significantly [2].

Artificial intelligence (AI) is becoming more prevalent in scientific research, speeding up discovery processes by assisting in hypothesis generation, experiment design, and data interpretation. This technology unlocks insights from large datasets, complementing traditional methods. Recent breakthroughs include self-supervised learning, allowing models to train on unlabeled data, and geometric deep learning, which improves accuracy and efficiency by utilizing scientific data structures. Generative AI techniques can produce designs such as small-molecule drugs and proteins by analyzing various data types, including images and sequences [3]. The results of medical diagnostics tests aid healthcare providers in determining optimal treatment plans for patients. Besides diagnosing conditions, diagnostics can monitor progress, assess treatment effectiveness, and detect potential health issues early. With the AI revolution, there's potential to enhance medical diagnostics by boosting prediction accuracy, speed, and efficiency. AI algorithms can analyze various medical data types, including 2D/3D imaging, bio-

signals (e.g., ECG, EEG), vital signs (e.g., body temperature), and medical history. This enables quicker and more accurate disease identification and diagnosis. Multimodal data analysis offers a smart solution, leveraging diverse patient data for better diagnostic decisions [4].

Machine learning is a rapidly growing field bridging computer science and statistics, integral to artificial intelligence (AI) and data science. Its progress is driven by new algorithms, theories, and abundant online data, alongside affordable computation. Developers recognize the efficiency of training systems with input-output examples over manual programming, given the multitude of potential inputs for various applications [5]. Machine learning technology is widely used in medical data analysis, particularly for focused medical diagnosis. It involves inputting existing medical records with known diagnoses into a computer program to run a learning algorithm. This process derives diagnostic knowledge from past cases, creating a classifier to aid physicians in diagnosing new patients, thus improving diagnostic speed, accuracy, and reliability. Additionally, it can train non-specialist students or physicians in diagnosing specific medical issues [6].

Image enhancement is a crucial technology in image processing, aiming to enhance image quality for applications. The fundamental principle involves adjusting the information content of an image to better suit a specific application [7]. In a study, the authors present MedGA, a new image enhancement technique utilizing Genetic Algorithms. MedGA aims to enhance images with a bimodal gray level intensity histogram by strengthening the two underlying sub-distributions. This method serves as a pre-processing step for improving the visual quality of such images and aiding anomaly detection and diagnosis [8].

Deep learning involves computational models with multiple layers that learn representations of data at different levels of abstraction. This approach has significantly advanced fields like speech recognition, visual object recognition, and object detection, as well as domains like drug discovery and genomics. By using the backpropagation algorithm, deep learning uncovers complex structures in large datasets by adjusting internal parameters to compute representations in each layer based on those from the previous layer [9]. Deep learning is increasingly applied in medicine for tasks like segmentation, diagnosis, classification, prediction, and detection of anatomical regions of interest. Compared to traditional machine learning, deep learning's ability to learn from raw data and extract complex abstractions from multiple hidden layers makes it superior. Its capability to utilize unlabeled data during training enables it to handle heterogeneous information effectively. While concerns exist about potential misuse, the positive applications of deep learning in medicine are extensive and promising [10].

This paper presents a new forensic method that utilizes convolutional neural networks (CNNs) to tackle two issues in manipulation detection. While existing CNNs excel at learning image content features, they face difficulty in capturing manipulation detection features. To address this, the paper introduces a novel CNN layer called a constrained convolutional layer. This layer not only suppresses image content but also adaptively learns manipulation detection features, offering a promising solution to enhance forensic analysis [11]. The paper delves into the utilization of Convolutional Neural Networks (CNNs), a prominent deep learning technique, for classifying pneumonia within a chest X-ray dataset. CNNs have demonstrated significant success in diverse image classification tasks, particularly in the realm of medical imaging. Despite their efficacy, the collection of medical image datasets remains challenging due to the expertise required for accurate labeling. The study seeks to assess the effectiveness of CNN-based algorithms in classifying pneumonia using chest X-ray images. [12].

This paper outlines the broad utility of Natural Language Processing (NLP) across multiple domains, including machine translation, email spam detection, information extraction, summarization, medical applications, and question answering. It presents four phases delineating different levels of NLP and components of Natural Language Generation. The historical progression and evolution of NLP are discussed, followed by a comprehensive examination of its current state of the art, highlighting various applications, ongoing trends, and challenges [13]. The research evaluates the effectiveness of a Natural Language Processing (NLP) tool for extracting medical issues from clinical documents in patients' electronic medical records. These issues are then recommended for inclusion in the patient's electronic problem list to improve its accuracy, comprehensiveness, and currency. The NLP system utilizes the UMLS MetaMap Transfer (MMTx) application and a negation detection algorithm called NegEx to identify 80 specific medical problems based on their frequency within the institution [14].

Transformers, the prevailing architecture in natural language processing, have garnered considerable interest from computational visual media researchers as well. This interest stems from their ability to handle long-range representations effectively and achieve high performance. Unlike traditional recurrent neural network (RNN) structures, transformers are sequence-to-sequence models that rely on a self-attention mechanism [15]. The paper discusses the effective integration of transformers in diverse clinical applications within medical image analysis, focusing on tasks like image synthesis, registration, segmentation, detection, and diagnosis. It starts with an introduction to the attention mechanism, a key aspect in transformers, and other fundamental components. Various transformer architectures tailored for medical image tasks are reviewed, alongside their limitations [16].

GANs, short for Generative Adversarial Networks, employ two neural networks in a zero-sum game, yielding more defined outputs. They are applied in image processing, video generation, and computer vision prediction. GANs work by setting a Generator against a Discriminator to produce synthetic data resembling real data. The Generator aims to capture the true data distribution to create new samples, while the Discriminator acts as a binary classifier to differentiate between actual and generated samples [17]. GANs are widely applicable across diverse tasks like synthetic data generation, image-to-image translation, segmentation, denoising, and super-resolution. They find extensive use in various supervised tasks, including disease classification, and organ, vessel, tumor, and region-of-interest segmentation. Specifically, a GAN architecture has been developed for segmenting lungs from original chest x-ray images, utilizing key components like convolution blocks, residual blocks, downsampling blocks, and upsampling blocks [18].

LLM represent artificial intelligence (AI) tools based on multi-layer recurrent neural networks that are trained on vast amounts of data to generate human-like text [19]. This study explores the potential of Large Language Models (LLMs) like OpenAI ChatGPT in medical imaging. LLMs, with their advanced language capabilities, aid radiologists in interpreting images, improve patient-physician communication, and optimize clinical workflows in hospital settings. Beyond clinical applications, these advancements impact various stakeholders within the medical imaging ecosystem [20].

Advancements in computing science and web technology offer the environmental community abundant resources for data collection and analysis, presenting new challenges in designing analysis methods, workflows, and data set interaction [21]. The rise of personal electronic devices and the internet has led to new communication platforms, offering patients access to information through various channels like chat rooms, social media, and blogs. Despite the abundance of medical information, low health literacy persists, leading to poorer health outcomes and higher disease rates [22].

Implementing an electronic health record (EHR) system is a complex process that requires careful planning to minimize errors. This involves evaluating selection criteria and creating an implementation plan focused on interoperability, confidentiality, availability, and integrity of patient health information data. It's crucial to ensure timely, accurate, and regulatory compliant generation of reports. The process includes assessing existing workflows in each department and determining the institution's requirements and preferences for the EHR system to function effectively [23].

Advancements in ultra-fast imaging have expanded the applications of magnetic resonance imaging (MRI) beyond diagnosis to therapeutic uses. With the integration of miniature intravascular imaging catheters and MRI-compatible devices, there are new opportunities in detecting and treating cardiovascular diseases. Early applications involve high-resolution intravascular and intracardiac imaging, along with procedures like balloon angioplasty, stent placement, and radiofrequency ablation. The potential extensions include therapeutic catheterization for pediatric cardiac cases. [24].

Lung imaging, notably using computed tomography (CT), is increasingly utilized alongside spirometry for comprehensive assessment of chronic obstructive pulmonary disease (COPD). While spirometry remains primary for diagnosis and severity staging, CT imaging has become crucial in clinical practice and research. COPD is diverse, with varying clinical features, radiographic presentations, progression rates, and outcomes. Recent studies underscore CT imaging's significant role in enhancing diagnostic accuracy, phenotyping, predicting disease progression and prognosis, selecting patients for interventions, and advancing understanding of COPD's complex pathophysiology. Several CT metrics show promise as valuable imaging biomarkers for COPD [25].

Older individuals taking numerous medications make up a significant and increasing portion of the population. Managing multiple medications can be difficult, particularly for seniors who often have multiple health conditions and physical or cognitive limitations. It's crucial for older adults to maintain good medication-taking abilities and adhere to their medication regimens to ensure safe and effective use of medications [26].

The focus of this study is on enhancing the usability and developer experience of machine learning tools for creative software developers and music technologists. The research reviews existing work on the design and evaluation of application programming interfaces (APIs) in the context of machine learning for music technology software development. The study introduces the RAPID-MIX API, designed to facilitate rapid prototyping with interactive machine learning, and presents findings from a usability evaluation involving software developers in the music technology domain [27].

3. Related Work

There are many works that are performed under the technologies that have been discussed. For instance, A group of researchers have made a comprehensive survey on brain tumour detection and its classification with the existing models discussing accuracies of the models taken and their limitations [28] Which is shown in Table 1. Next, a study developed a 22-layered convolutional neural network (CNN) structure to classify brain tumour types using 3064 T1-weighted contrast-enhanced MRI images. Their model successfully categorized

meningioma, glioma, and pituitary tumours with an accuracy of 96.56% [29]. A study proposed 2D CNN achieved a training accuracy of 96.47%, while the proposed auto-encoder network achieved a slightly lower accuracy of 95.63%. The average recall values for the 2D CNN and auto-encoder networks were 95% and 94%, respectively. They also compared machine learning methods like, Multilayer Perceptron (MLP) which had lowest accuracy rate of 28%, while K-Nearest Neighbours (KNN) achieved the highest accuracy rate of 86%. [30]. Then, a study proposed a framework for tumour classification based on MRI data that combines deep neural networks with a novel data distillation and augmentation procedure to combat scarcity of the training data and labelling noise which analyses the medical images like a professional medical practitioner [31]. Another study proposed several methods for segmenting brain tumours using datasets from BraTS 2019, 2020, and 2021, which consist of multi-modality brain MRIs comprising various slices. Previous research has investigated whether combining 2D slices improves accuracy compared to using 3D MRI data [32]. This document describes a brain tumor segmentation system that utilizes preprocessing to focus on specific image sections, reducing computational time and overfitting. It employs a Cascade Convolutional Neural Network (C-ConvNet/C-CNN) model that explores both local and global features. A Distance-Wise Attention (DWA) mechanism enhances model accuracy. Experiments on the BRATS 2018 dataset show competitive results, with mean whole tumor, enhancing tumor, and tumor core dice scores of 0.9203, 0.9113, and 0.8726, respectively [33].

This paper examines the use of artificial neural networks, a key component of deep learning, for analyzing large datasets to identify patterns and forecast outcomes. It discusses the growing adoption of deep learning in cancer biology analysis, driven by its efficacy and the availability of cancer-related data. Specifically, it explores advanced deep learning methods and their application in analyzing transcriptome, methylation, and genomic data in cancer research. The paper also addresses challenges in applying deep learning to precision oncology, including the need for interpretable models and limited phenotypic data. It concludes by suggesting strategies to overcome these challenges for future clinical use [34].

The above-mentioned models mainly discuss on brain tumour classification, and when it comes to brain MR Image analysis not only tumour but also Alzheimer's can be evaluated using brain MRI. A study showcases a classification method utilizing Random Forest (RF) feature selection combined with CNN classification. The model is trained using a diverse group of patients including those with healthy cognition (HC), Alzheimer's disease (AD), mild cognitive impairment (MCI), and cognitive MCI (cMCI). The study achieved an overall accuracy rate of 38.8% [35]. Recent research has highlighted the challenge of diagnosing Alzheimer's disease (AD) in rural areas with limited clinical resources. Machine learning methods, particularly deep learning, have shown promise in using MRI scans to distinguish between normal cognition, mild cognitive impairment (MCI), AD, and non-

Alzheimer's dementia. A recent study developed a deep learning framework achieving accuracy comparable to diagnoses by neurologists and neuroradiologists, across these cognitive states. This framework, incorporating MRI data and other variables, demonstrates deep learning's potential in improving cognitive disorder diagnosis and understanding neurodegenerative markers [36].

A study utilized structural MRI and machine learning to enhance brain tumor classification and diagnose Alzheimer's disease (AD) and mild cognitive impairment (MCI). Employing data from OASIS and ADNI databases, classifiers achieved a balanced accuracy of 90.6% and MCC of 0.811 for "HC vs. AD" and 62.1% for "HC vs. MCI vs. AD". Hippocampal traits were most influential in classification, followed by temporal, cingulate, and frontal areas. Graph theory measures did not improve performance. The proposed tool, trained on baseline scans and follow-up diagnosis, offers adaptability across datasets and methods. [37].

Lung cancer is a leading cause of death, often detected in early stages as pulmonary lung nodules. Deep learning algorithms, including state-of-the-art object detection methods, aim to enhance nodule detection accuracy. Despite challenges in obtaining clinically relevant results, a systematic comparison of these algorithms highlights strategies such as data augmentation and transfer learning to address class imbalance and improve performance. The model, which won the Node21 competition's detection track, showcases cutting-edge performance in lung nodule detection [38]. Another study indicates that the effectiveness of the CAD system in diagnosing lung disorders relies on accurate lung segmentation throughout the multi-stage process of automatically evaluating lung CT images using deep-learning [39].

Furthermore, these methodologies are also useful in covid-19 and pneumonia detection. A proposed COVID+CAP-CNN framework achieved slice-level classification accuracy of over 94% and validation accuracy of 89.3% at identifying COVID-19 and CAP (Community Acquired Pneumonia) [40]. Apart from the above research there are additional advancements and contributions in AI, ML, and NLP from Flatiron Health, including model-assisted cohort selection [41], prediction of future health events [42], continuous bias monitoring software [43], automated mapping of laboratory data [44], and point-of-care products aimed at enhancing patient care and clinical trials [45]. The study utilizes computed tomography images and a convolutional neural network (CNN) for COVID-19 testing. It concludes that employing larger out-of-field datasets improves the testing capability of pre-trained models. The suggested transfer learning strategy shows superior identification performance compared to current methods. Evaluation on randomly chosen training datasets indicates satisfactory performance. The model's visual features may aid manual screening by clinical physicians [46].

Next advancement in the medical field is the examination of the EHR (Electronic Health Report) using Natural Language Processing. Which is further discussed in the methodology section.

A study examined deep learning, machine learning, and artificial intelligence models used in medical image analysis, noting challenges arising from dataset distortions during transfer and saving processes. Image enhancement methods, like resolution enhancement and noise reduction, were explored to mitigate these challenges and improve image quality. The study highlighted the importance of enhancing image resolution to enhance the accuracy of medical image analysis [47].

Research was proposed where they use Image Enhancement techniques on Brain MRI for tumour classification which consists of gaussian-blur-based sharpening and adaptive histogram equalization using CLAHE Contrast-Limited Adaptive Histogram Equalization which reportedly increased the accuracy when compared to the models which did not use image enhancement techniques [48]. Another study has concentrated on Image Enhancement techniques on classification of Alzheimer's disease which was performed on AD using 2D and 3D brain images. From their three-class dataset, they extracted features using Gray Level Co-Occurrence Matrix and Gabor Filter techniques. Features were selected with voxel-based morphometry method and with the help of the SVM classifier they achieved a classification success of 92.86% [49].

Similarly, like the above techniques which were performed on Brain MRI scans Image enhancement techniques were also performed on Chest CT Scans. A study has collected small cell lung cancer CT scan dataset images and these images are then examined by applying multilevel brightness-preserving approach which effectively examines each pixel, eliminates the noise, and increase the quality of the lung image [50]. Another study used a publicly available chest x-ray image dataset, as the image quality will be degraded while taking x-rays, they have applied image quality enhancement using median filtering followed by histogram equalisation [51].

Now, the above-mentioned models were basically discussed on MRI Scans and CT scans. But when it comes to enhance the images of the report that are captured by the user will get degraded due to technical issues and many environmental impacts during the image acquisition, these issues can be resolved by a proposed methodology of increasing the quality of the Image using Restoration which is instructed by humans [52].

There are some models which focused on report generation from the scan images input taken from database. One of those models introduce a new framework called PromptMRG, which stands for diagnosis-driven prompts for medical report generation. This innovative approach aims to enhance the diagnostic precision of

medical report generation by incorporating prompts that are sensitive to the diagnosis [53]. Another model was focused on Medical Vision-Language (Med-VL) encompass Visual Question Answering (Med-VQA), medical report generation, and medical image-text retrieval. These tasks have garnered increasing attention due to their potential in computer-assisted diagnosis and healthcare automation [54].

There is a study outlines specific objectives or aims, which may involve identifying risk factors for cognitive decline, understanding the connections between inflammation and cognitive function, or suggesting interventions to prevent or reduce cognitive decline in older adults from Parkinson's disease [55]. A study introduces a robust framework utilizing machine learning to predict long-term stroke risk, incorporating a stacking technique. This approach achieves exceptional performance metrics: accuracy of 98%, recall of 97.4%, precision, F-measure, and AUC of 98.9%. Outperforming other methods, early stroke symptom identification enhances prevention efforts and encourages healthier lifestyles [56].

Generative AI involves algorithms and models generating new content, like images or text, rather than just recognizing patterns. It aims to produce output resembling human-created content. In drug discovery, AI integration speeds up the process by using computational techniques to identify potential drug candidates for treating diseases [57]. Deep learning-based generative models are emerging as a promising approach for de novo design of molecules, offering a potentially faster and more efficient alternative to traditional drug discovery methods [58].

Digital Twins (DTs) are emerging as powerful tools in drug discovery and clinical trials. These virtual representations of biological systems, ranging from individual cells to entire patients, enable scientists to conduct in silico (computer-based) simulations and experiments [59]. Generative AI models hold immense potential for transforming various aspects of healthcare. They can generate human-like text, images, and even molecules, offering: Enhanced clinical decision support: AI models can analyse medical records, suggest diagnoses, and personalize treatment plans [60]. AI and machine learning (ML) are rapidly transforming drug discovery. This raises questions about the impact on clinical pharmacology, the field responsible for understanding drug effects in humans [61].

Traditional drug discovery is slow, expensive, and often leads to dead ends. Finding promising drug candidates among vast chemical spaces is a significant hurdle [62]. Generative AI is revolutionizing drug discovery: By generating novel molecules with desired properties, it offers a faster and more efficient alternative to traditional methods [63]. AI is transforming drug discovery by: Optimizing lead compounds based on various properties like efficacy, safety, and drug-likeness. Identifying potential drug targets and repurposing existing drugs [64].

Generative AI offers exciting possibilities for laboratory medicine: Automating tasks: Generate reports, write clinical notes, and design experiments, Personalized medicine: Develop tailored diagnostic tools and treatment plans, Drug discovery: Identify new drug targets and design potential therapeutics [65]. ChatGPT, a large language model (LLM), represents a new wave of generative AI with potential applications in science and healthcare, Potential benefits include: Generating scientific text, reports, and grant proposals: Accelerating research and communication [66].

Now, the above-mentioned models have for sure performed well in their respective fields by acquiring high accuracies. Similarly for Gen-AI and Large Language models many were introduced. But our model has evaluated all the other existing models on the respective fields and achieve higher accuracy than the others, making our model more efficient and with the infusion of most accurate Large Language Model i.e., GPT-4. This made our model more reliable for the public. Our model is further evaluated in the proposed framework section, methodology section, and discussion section.

4. Proposed Framework

The proposed model is comprised of many technologies like web technologies for the generation of chat interface for users, image pre-processing methods like image enhancement, de-noising, de-hazing, CLAHE, classification of tumour, Alzheimer, covid-19, using deep learning, image classification like CNN, ViT, machine learning, artificial intelligence, EHR analysing using NLP, prompts using LLMs, API references.

For a user to access this framework he/she must have an internet connection and should have access to any web browser to search and utilise it.

The user interface of the framework is built using web technologies where the user can chat with the model for which we have included features like audio transcription, and image uploading. This interface takes the image of brain scan or lung scan or medical report or query from the user. Processing the query from the user the backend comes into play.

Now, the backend is comprised of image pre-processing techniques like image enhancement, image de-noising, image de-hazing, low-light and CLAHE (Contrast-Limited Adaptive Histogram Equalization). These techniques are proven to improve the overall quality of the MRI Scan Images, CT Scan Images, and medical reports captured by user using Image Capturing Devices.

Once the above pre-processing work is done the output image is then forwarded to image classification techniques which is comprised of brain tumour

classification, Alzheimer's disease classification, lung tumour classification, covid-19 classification, along with those medical reports are also classified.

Once the classification is done, each of these methods undergoes evaluation process to predict the possible outcomes that are to be generated from the input images after image pre-processing techniques are also evaluated generating output for the classification methods. The images are then sent for evaluation of every classification models. Which are discussed in a detailed manner in methodology and discussion sections of this paper.

After the techniques, the possible findings are generated in a form of report and once the report is generated these are evaluated with the possible medical prompts and are analysed using large language models (LLMs) and output the possible and accurate report, and once the report is generated it is then displayed onto the user interface for the user to examine and understand.

This does not stop here after the possible report is in front of the user it is basic human nature that doubts will arise. Once a user gets doubt, they can post that query in our model and once it is posted the ChatGPT API Key reference comes into play which takes the input from the user and then it is evaluated with all the possible medical prompts that are ingested in the model and then generate most accurate result for the query.

This framework may sound complicated and for that we have designed the framework, which is displayed as figure 1.

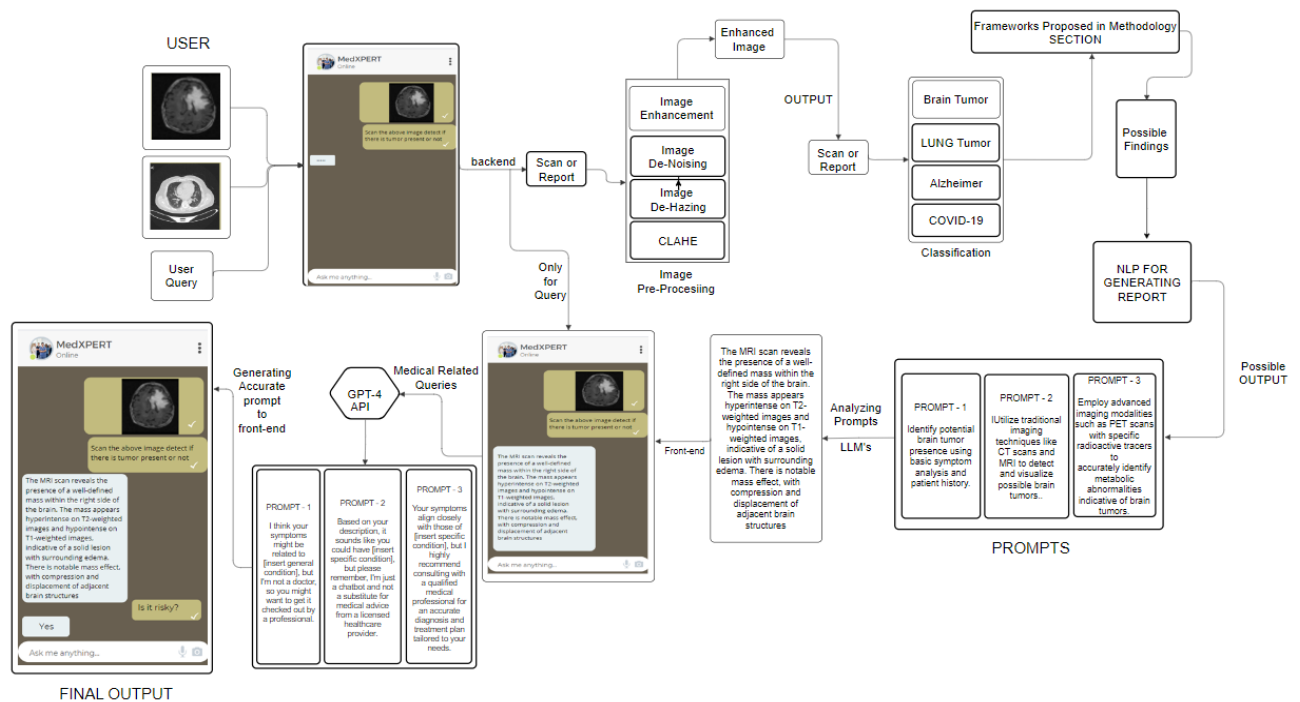


Figure 1 Proposed Framework of MedXPert

The above framework is the skeleton of the model MedXPert. Which is a combination different models, prompts, and API reference making this is a reliable, accessible to the public, and efficient of the existing or referred models.

5. Methodology for Image Pre-Processing and Medical Diagnosis

5.1 Brain Tumour Detection

Brain tumors, resulting from the uncontrolled growth of brain cells, can adversely affect health by exerting pressure on the brain within the skull. Early detection and classification are crucial in selecting appropriate treatment methods. They are commonly categorized as benign or malignant. Benign tumors, like meningiomas, can still pose risks and may rarely transform into malignancies but are often surgically removable.

As per the World Health Organization (WHO), brain tumors are classified into grades I-IV. Tumors categorized as grades I and II are characterized by slow growth, while grades III and IV tumors are more aggressive and associated with a less favourable prognosis [67].

Grade I tumors grow slowly and do not spread quickly, offering better long-term survival chances and often being fully removable through surgery, as seen with pilocytic astrocytoma. Grade II tumors also grow slowly but can infiltrate nearby tissues and potentially recur after surgery, exemplified by oligodendroglioma. Grade III tumors progress faster than Grade II, invading neighbouring tissues, requiring additional treatment like radiotherapy or chemotherapy post-surgery, as observed in anaplastic astrocytoma. Grade IV tumors are the most aggressive, spreading rapidly, possibly utilizing blood vessels for growth, such as glioblastoma multiforme [28].

The Brain can be scanned by various studies on techniques. One is PET tracers which lack precision in pinpointing anatomical structures due to their inferior spatial resolution compared to MRI scans [68]. Other MRI can assess structures that may be difficult to discern in a CT scan, offering superior contrast for soft tissues and enhancing the clarity of anatomical structures [69]. And An MRI scan is utilized for comprehensive analysis of various body parts and is particularly effective in identifying brain abnormalities at earlier stages compared to other imaging techniques [70].

There are multiple performance measures when it comes to the evaluation of brain tumor. They are Peak Signal to Noise Ratio, Jaccard Index, Structured Similarity Index, Dice Index, Mean Squared Error, Area Under Curve, Sensitivity, Accuracy, Positive Predictive Value, and Specificity [28].

Without Image Pre-Processing:

For this the datasets that we have used are:

1. The first dataset, named the Reference Image Database to Evaluate Therapy Response (RIDER), comprises MRI multi-sequence images from 19 patients diagnosed with glioblastoma (Grade IV). This dataset contains a total of 70,220 images.
2. The second dataset, called The Repository of Molecular Brain Neoplasia Data (REMBRANDT), consists of MRI multi-sequence images from 130 patients with gliomas of Grade II, Grade III, and Grade IV. The total number of images in this dataset is 110,020.
3. The third dataset is known as The Cancer Genome Atlas Low-Grade Glioma (TCGA-LGG), which contains 241,183 MRI images of 199 patients diagnosed with low-grade gliomas (Grade I and Grade II).
4. The fourth dataset used in the study contains 3064 T1-weighted contrast-enhanced images from 233 patients with three types of brain tumors: glioma (1426 slices), meningioma (708 slices), and pituitary tumors (930 slices).

Each dataset includes MRI images with T1-contrast-enhanced and FLAIR sequences, providing multimodal information for analysis. These datasets are part of the Cancer Imaging Archive (TCIA) project.

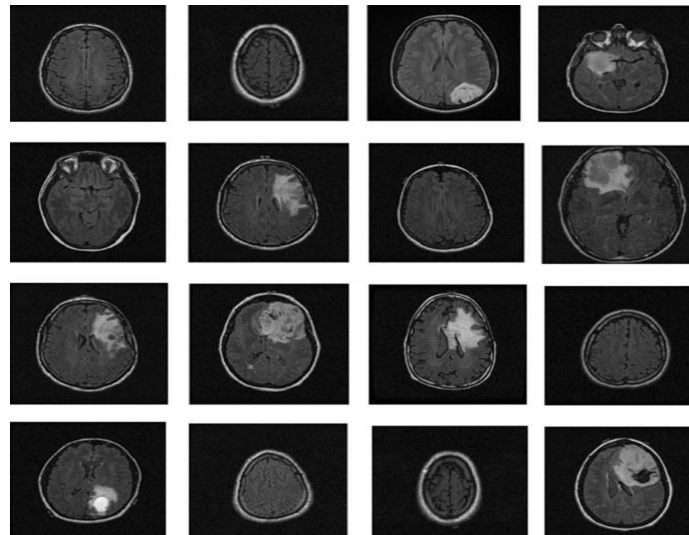


Figure 2 Reference of the dataset used with different grades

Classification -1 method:

For Classification-1 task, a total of 2990 images are collected, including 1640 tumor and 1350 no tumor images.

The CNN model proposed for Classification-1 comprises 13 weighted layers, including 1 input layer, 2 convolutional layers, 2 ReLU layers, 1 normalization

layer, 2 max pooling layers, 2 fully connected layers, 1 dropout layer, 1 softmax layer, and 1 classification layer. The model is designed to classify images into 2 classes, reflected in the output layer having two neurons. The last fully connected layer produces a two-dimensional feature vector, serving as input to the softmax classifier, which determines the final prediction regarding the presence of a tumor.

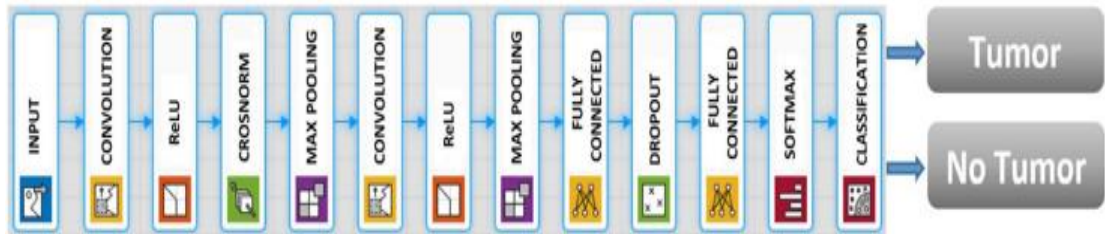


Figure 3 Architecture of the proposed CNN model for Classification-1 task

Classification -2 method:

For Classification-2 task, a total of 3950 images are collected, including 850 normal, 950 glioma, 700 meningioma, 700 pituitary and 750 metastatic images.

The proposed CNN model for Classification-2 consists of 25 weighted layers, including 6 convolutional layers, 6 ReLU layers, 1 normalization layer, 6 max pooling layers, 2 fully connected layers, 1 dropout layer, 1 softmax layer, and 1 classification layer. The output layer has five neurons since the model is designed to classify images into 5 classes representing different tumor types. The last fully connected layer produces a five-dimensional feature vector, which is then fed into a softmax classifier to make the final prediction about the tumor type.



Figure 4 Architecture of the proposed CNN model for Classification-2 task

Classification -3 method:

For Classification-3 task, a total of 4570 images are collected, including 1676 grade II, 1218 grade III and 1676 grade III.

The proposed CNN model for Classification-3 consists of 16 weighted layers, including 3 convolutional layers, 3 ReLU layers, 1 normalization layer, 3 max pooling layers, 2 fully connected layers, 1 dropout layer, 1 softmax layer, and 1 classification layer. The output layer contains three neurons, corresponding to the three classes for tumor grade classification. The final prediction about the tumor grade is made by the softmax classifier, which takes the three-dimensional feature vector from the last fully connected layer as input.

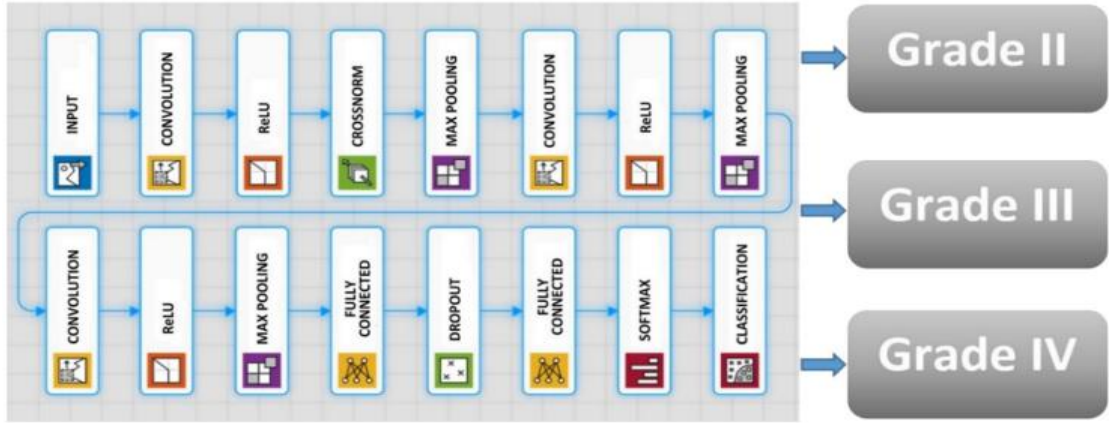


Figure 5 Architecture of the proposed CNN model for Classification-3 task

The increasing complexity of Convolutional Neural Networks (CNNs) in medical image processing presents challenges due to deeper architectures and higher-quality input images, leading to greater computational costs. To address this, powerful hardware and optimized hyperparameters are essential. Grid search optimization is commonly employed to automatically tune important hyperparameters, especially when the value range is relatively small. This method aims to find the best combination of hyperparameters by training the network across all specified range combinations. CNN models involve numerous hyperparameters, categorized as architectural (e.g., number of layers, filters, activation function) and fine adjustment (e.g., regularization, momentum, batch size, learning rate). In this study, architectural hyperparameters are tuned first using Algorithm 1, followed by fine adjustment hyperparameters using Algorithm 2 once the architectural hyperparameters are determined.

In the proposed study, grid search is performed on the training set using a fivefold cross-validation procedure. The dataset is divided into fivefold, with 2990 images for Classification-1, 3950 images for Classification-2, and 4570 images for Classification-3 tasks. For each classification task, the dataset is randomly separated into training, validation, and test sets in a 60:20:20 ratio. Algorithm 1 optimizes five parameters, each with varying numbers of combinations (4, 4, 7, 5, and 4). Hence, the total combinations checked are 2240. The grid search for the CNN model's architectural hyperparameters is executed 11,200 times due to the fivefold cross-validation procedure. Similarly,

Algorithm 2 optimizes four parameters with different combinations (4, 4, 5, and 4), resulting in 320 total combinations checked. The grid search for fine-tuning hyperparameters of the CNN model is executed 1600 times with fivefold cross-validation.

Algorithm 1: Grid Search algorithm to optimize the architectural hyper-parameters	
Step 1: Set 5-dimensional grid for five hyper-parameters to be optimized	: <i>Number of convolutional and max pooling layers</i> : <i>Number of FC layers</i> : <i>Number of filters</i> : <i>Filter sizes</i> : <i>Activation function</i>
Step 2: Set potential value intervals corresponding each dimension	: <i>Number of convolutional and max pooling layers</i> = [1, 2, 3, 4] : <i>Number of FC layers</i> = [1, 2, 3, 4] : <i>Number of filters</i> = [16, 24, 32, 48, 64, 96, 128] : <i>Filter sizes</i> = [3, 4, 5, 6, 7] : <i>Activation function</i> = [ELU, SELU, ReLU, Leaky ReLU]
Step 3: Look for all the candidate combinations and get the best one which optimizes the overall accuracy	: e.g. $C_1 = (2, 4, 16, 3, ELU) \rightarrow \text{acc} = 95\%$, $C_2 = (4, 4, 96, 7, SELU) \rightarrow \text{acc} = 97\%$, $C_3 = (2, 3, 64, 6, ReLU) \rightarrow \text{acc} = 99\%$
Algorithm 2: Grid Search algorithm to optimize the fine adjustment hyper-parameters	
Step 1: Set 4-dimensional grid for four hyper-parameters to be optimized	: ℓ_2 regularization : <i>Momentum</i> : <i>Mini-batch size</i> : <i>Learning rate</i>
Step 2: Set potential value intervals corresponding each dimension	: ℓ_2 regularization = [0.0001, 0.0005, 0.001, 0.005] : <i>Momentum</i> = [0.80, 0.85, 0.9, 0.95] : <i>Mini-batch size</i> = [4, 8, 16, 32, 64] : <i>Learning rate</i> = [0.0001, 0.0005, 0.001, 0.005]
Step 3: Look for all the candidate combinations and get the best one which optimizes the overall accuracy	: e.g. $C_1 = (0.0001, 8, 0.8, 0.0001) \rightarrow \text{acc} = 94\%$, $C_2 = (0.0005, 32, 0.9, 0.001) \rightarrow \text{acc} = 98\%$, $C_3 = (0.001, 64, 0.9, 0.001) \rightarrow \text{acc} = 99\%$

With Image Pre-Processing:

The above methodology is integrated with Gaussian-blur-based sharpening and Contrast-Limited Adaptive Histogram Equalization (CLAHE) with the proposed model to facilitate more precise diagnostic procedures for identifying glioma, meningioma, pituitary tumors, and cases without malignancies, including whether tumor is present or not, and detecting the Grade of it. This whole architecture is designed which is shown in figure 6.

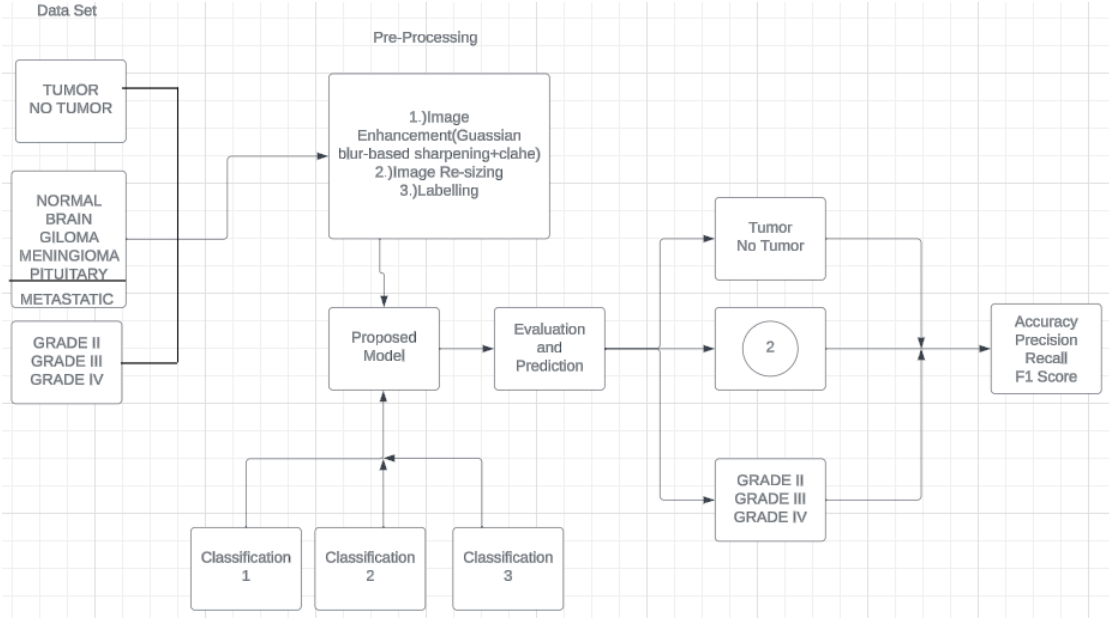


Figure 6 Proposed Framework for Brain Tumor detection using image pre-processing

We implemented a preprocessing framework to improve image quality by integrating sharpening and Contrast-Limited Adaptive Histogram Equalization (CLAHE) approaches. The process of sharpening commenced by implementing a Gaussian blur through the utilization of a specific technique. The utilization of a 5×5 kernel was suitable in the process of attenuating high-frequency noise. The resultant enhanced image was determined using the formula:

$$\text{Sharpened Image} = 1.5 \times \text{Original Image} - 0.5 \times \text{Blurred Image}$$

Subsequently, the image underwent a conversion process to grayscale, facilitating a precise enhancement of contrast. To achieve this, CLAHE was utilized, characterized by an 8×8 -tile grid and a clip limit of 2.0. Distinct from global histogram equalization, CLAHE adopts a localized strategy by partitioning the image into discrete tiles and performing individual equalizations, encapsulated by

$$H_{\text{local}}(i) = \text{CLAHE}(H_{\text{tile}}(i))$$

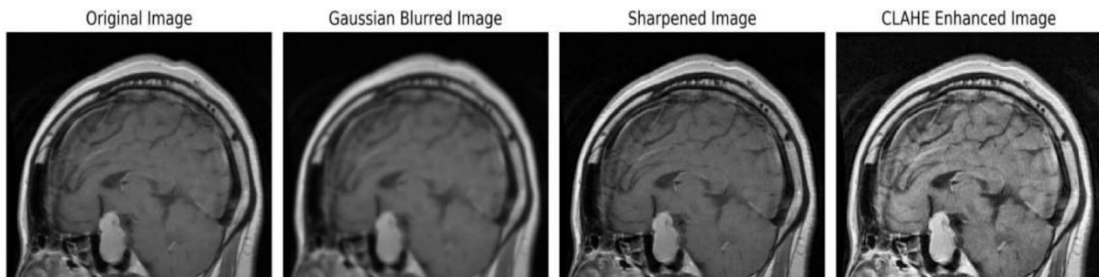


Figure 7 Sequential image improvement as part of the preprocessing framework. The stages progress from the unaltered original image through Gaussian blurring for noise suppression, sharpening the emphasized edge definition to the final enhancement using CLAHE.

The output of the model combining both the methods pre-processing and original model is present in the discussion section under the same brain tumor detection.

5.2 Alzheimer Disease Detection

Alzheimer's disease occurs because of the loss of cognitive functions in the brain causing near-forgetfulness and dementia in subsequent processes. Dementia is a disease that causes loss of talent during the action. The most common type of dementia in the world is Alzheimer's disease (AD) with a rate of about 80%.

Alzheimer's disease is a primary cause of dementia, particularly in older individuals, due to its gradual neurodegenerative process. The disease progresses through five stages: Subjective Memory Concern (SMC), Mild Cognitive Impairment (MCI), Early MCI (EMCI), Late MCI (LMCI), and finally Alzheimer's Disease (AD). Diagnosis of Alzheimer's typically involves a conventional MRI brain scan [71].

The data set used in this model is available on Kaggle has been personally verified by the uploader and consists of 6400 samples. Each sample is an individual three-channel (RGB) image with dimensions of 176 x 208 pixels. These samples are divided into four different classes, with the following distribution: 3200 samples belong to the NOD class, while the remaining three classes (VMD, MD, and MOD) contain 2240, 896, and 64 images, respectively. Additionally, the data set size is considered reasonable, and the images have already been cleaned, resized, and organized, making them suitable for research purposes.

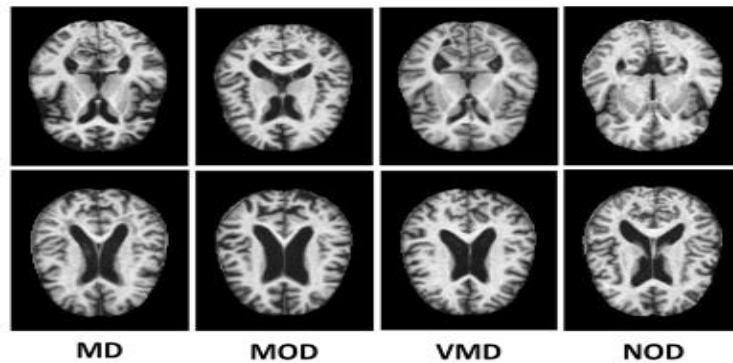


Figure 8 Image samples from AD dataset without up-sampling

Without Image pre-processing

The presented model effectively identifies early stages of Alzheimer's disease and displays class activation maps as heatmaps on the brain. The proposed Alzheimer's Disease Detection Network (ADD-Net) is developed from scratch

to accurately classify AD stages while reducing parameters and computational costs.

In this proposed model, the input dataset undergoes normalization as a preprocessing step. Categorical data variables are converted using the one-hot encoder before being fed into the ADD-Net. To address the issue of imbalanced

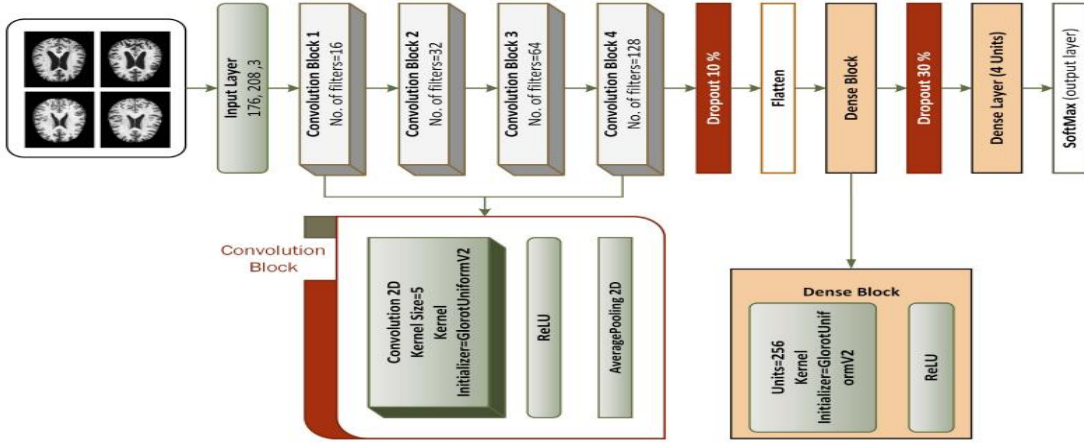


Figure 9 Architecture of the proposed ADD-Net for early detection of AD.

data, the Synthetic Minority Oversampling Technique (SMOTETOMEK) algorithm is applied to oversample the classes and balance the dataset. Subsequently, the dataset is split into training, testing, and validation sets in a ratio of 60%, 20%, and 20%, respectively. Feature extraction is performed using a standard CNN to effectively train the ADD-Net, as depicted in Figure 9. The size of training parameters is intentionally kept smaller than in previous studies for enhanced robustness in Alzheimer's disease (AD) classification. Additionally, the Grad-CAM heat-map algorithm is employed to visualize the class activation map, highlighting the features contributing to the classification of each image sample.

The Convolutional Neural Network (CNN) architecture mimics the structure of the human brain and is primarily utilized in computer vision tasks like image classification, segmentation, and object detection. Its translation-invariant nature has made it a preferred choice for deep learning models. This paper introduces a new CNN model, ADD-Net, specifically designed for accurate Alzheimer's disease (AD) classification. ADD-Net consists of four convolutional blocks, each incorporating a Rectified Linear Unit (ReLU) activation function, a 2D average pooling layer, two dropout layers, two dense layers, and a SoftMax classification layer, as illustrated in Figure 9. The detailed network architecture and model summary, along with the hyper-parameter descriptions essential for practical training of ADD-Net, are provided in figure 10 below.

Model Summary		
Layer Type	Output Shape	Parameters
Input Layer	(None, 176, 208, 3)	0
ADD-NET Block01	(None, 86, 102, 16)	1216
ADD-NET Block02	(None, 41, 49, 32)	12832
ADD-NET Block03	(None, 18, 22, 64)	51264
ADD-NET Block04	(None, 7, 9, 128)	204928
Dropout_1	(None, 7, 9, 128)	0
Flatten	(None, 8064)	0
Dense_1	(None, 256)	2064640
Dropout_2	(None, 256)	0
Dense_2	(None, 4)	1028
Output: SOFTMAX	(None, 4)	0
Total Parameters		23,35,908
Trainable Parameters		23,35,909
Non-Trainable Parameters		0

Figure 10 Total parameters for the proposed ADD-Net model.

With Image pre-processing

The dataset is re-enhanced separately with DeepDream, Fuzzy Color Image Enhancement, and hypercolumn techniques. VGG-16 (Visual Geometry Group-16) is used in the enhancing process in which deep features is also combined and linear regression is used for the selection of efficient features and support vector machine is preferred as a classifier.

The **DeepDream** is a simulation technique that is based on the imaginary dimensions of the human brain, using the features of the input images. The DeepDream is developed by Google, and the open-source code of the DeepDream is supported using the TensorFlow library. The DeepDream aims to find patterns in images through algorithmic pareidolia and to develop the patterns found using a CNN. That is, it makes the patterns (feature vector carrying the label of an object) seen in a particular image appear on the data by processing them with pretrained data. As a result, the DeepDream algorithm is useful for rendering visual content in surreal and abstract styles in over-processed images.

The working logic of the DeepDream algorithm is as follows; when an image is an input to a trained neural network model, neurons fire and activations occur. The DeepDream algorithm selects some of these neurons, allowing them to fire more than others. So, it increases their activation. The activation boost is accomplished by gradient ascent. This process is repeated until it contains all the features that the layer was originally looking for.

In **FCIE (Fuzzy Color Image Enhancement) Algorithm** plays an important role in image analysis. This algorithm consists of three stages function. These stages are as follows; image coding, fuzzy enhancement, and image decoding. Thanks

to these functions, the gray level intensities in each image are matched to a fuzzy plane. In the first stage, each image in the dataset is converted from gray level area to fuzzy level area. At this stage, membership value is given for each pixel. That is, each gray level pixel is assigned a membership degree depending on its position in the histogram.

In the second stage, the fuzzy level parameters used to change the image are updated and the aim is to enhance the fuzzy in the image. In the last stage, fuzzy enhanced images are decoded to re-convert them into gray-level images. Each pixel in the fuzzy level is converted into gray-level pixels according to the membership degree.

Calculation of the membership value the formulas between Eqs. (1) and (3) are used. In these equations; F_d and F_e are conversion coefficients, f_{max} denotes the maximal gray value, f_{ij} denotes the gray level of the (i, j) th pixel, $T^{(r)}$ is defined as successive applications of T .

$$\mu_{ij} = \left[1 + \frac{f_{max} - f_{ij}}{F_d} \right]^{-F_e} \quad (1)$$

$$\mu_{ij} = T^{(r)}(\mu_{ij}); r = 1, 2, 3, \dots, \quad (2)$$

$$T(\mu_{ij}) = \begin{cases} 2(\mu_j)^2; & 0 \leq \mu_j \leq 0.5 \\ 1 - 2(1 - \mu_j)^2; & 0.5 \leq \mu_j \leq 1 \end{cases} \quad (3)$$

Hypercolumn is a technique that performs the classification of pixels using hypercolumn. That is, each image given as an input to the model has a hypercolumn vector. These hyper vectors hold all the activation features of that pixel in the convolutional model. Thus, instead of deciding according to the pixel value in the final layer of the convolutional model in the classification process, it chooses the most efficient one by examining all the features in the hypercolumn vector.

The essence of the Hypercolumn technique is based on heat maps. After the convolution layers of the model, this technique uses bilinear interpolation and creates a transition feature value using two feature values with Bilinear interpolation. In other words, bilinear interpolation creates a smooth transition value between two feature values. In this way, feature maps extracted from other layers of the model are added, and it is processed with the sigmoid function. Heat maps extracted from the model are then combined to produce possible output values. This joining is done by the ‘‘Concatenate’’ function in the Hypercolumn technique. The purpose of our use of the Hypercolumn technique is to obtain a new dataset by enhancing the original dataset in figure 8.

With the help of image pre-processing the images in the dataset are very well enhanced as shown in the figure 11 below.

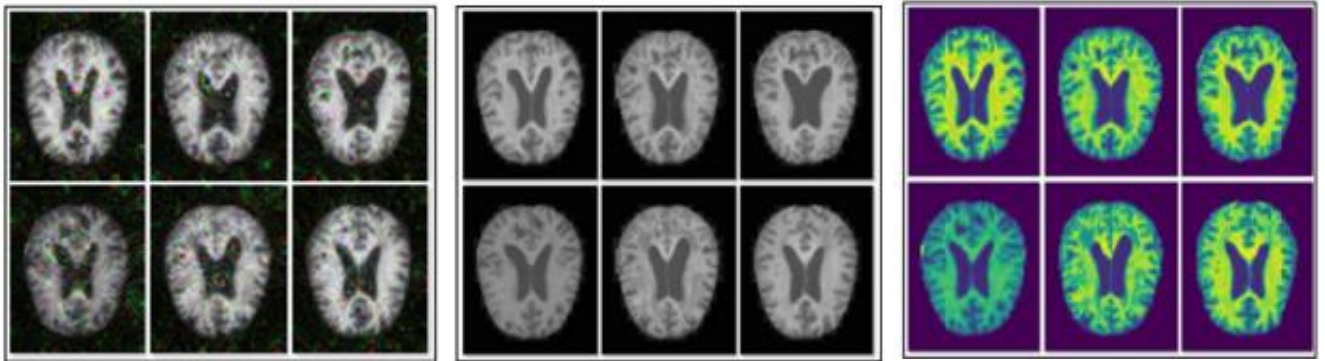


Figure 11 The three figures indicate that the images in the database have been pre-processed with the techniques mentioned above and are ready to get classified

The enhancement of the images and classification was achieved because of the below framework in figure 12.

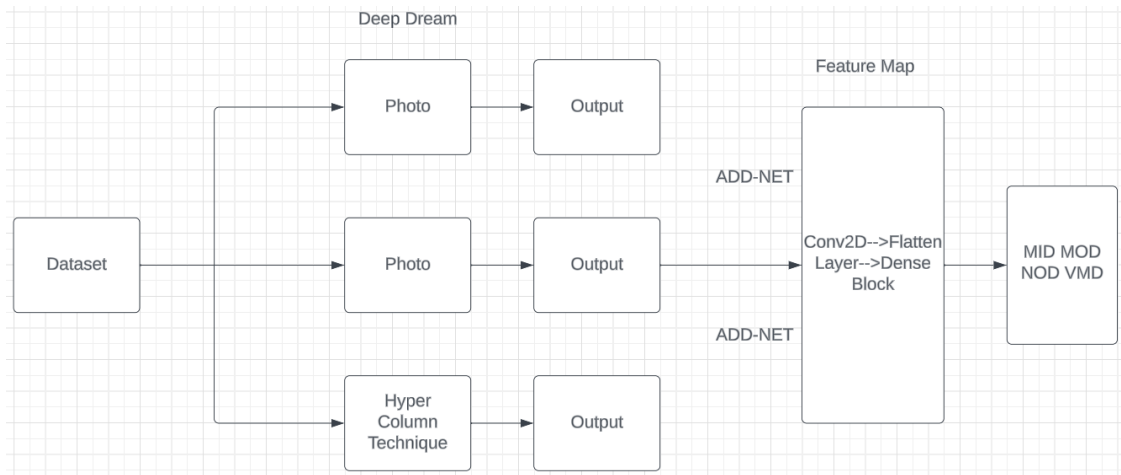


Figure 12 Proposed framework for the detection of Alzheimers using Image Pre-Processing techniques

The results of the model combining both the methods pre-processing and original model is present in the discussion section under the same Alzheimer detection.

5.3 Lung Tumor Detection

Among various disease examination processes lung cancer needs more attention as it effects both male and female, which also causes high mortality rate and one of the main complications of lung tumour detection is that it is not identified easily and once it gets identified the damage is already done [50]. While computed tomography (CT) scans are commonly used for diagnosing lung cancer, manual analysis is slow and susceptible to errors. To address these issues, computational methods, particularly machine learning and deep

learning algorithms, are being utilized to expedite and enhance the accuracy of identifying cancerous and non-cancerous CT scans.

Experiments were conducted using the lung cancer dataset from the Iraq-Oncology Teaching Hospital/National Centre for Cancer Diseases (IQ-OTH/NCCD), collected over a period of more than three months in 2019. This dataset includes CT scans from individuals diagnosed with various stages of lung cancer, annotated by multiple oncologists and radiologists. It consists of 1097 chest images representing CT scan slices from 110 cases, with variations in age, gender, education status, residence, and living status. Cases were categorized as benign, malignant, or normal, totalling 110 cases, including 40 malignant, 15 benign, and 55 normal cases. Initially in DICOM format, the CT scans were later converted to JPEG for easier accessibility. The IQ-OTH/NCCD dataset is publicly available on Kaggle.

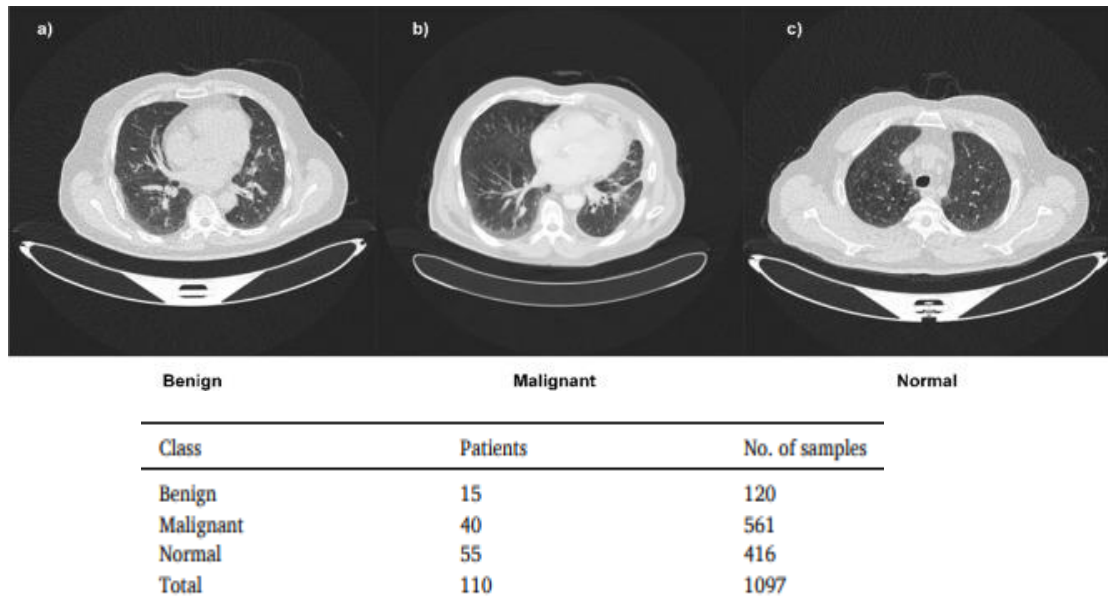


Figure 13 The above figure depicts the images in the Kaggle database that is considered, the below depicts the patients and no. of samples in the same database

Without Image pre-processing

This model is for classifying lung cancer, employing a transfer learning approach. This is constructed by adapting the architecture of EfficientNet and enhancing it with additional layers in the classification head. The evaluation of this model involves the use of five different variants of EfficientNet, namely, B0 to B4.

Transfer Learning, in contrast to traditional machine learning methods, Convolutional Neural Networks (CNNs) enable automatic extraction of both low-level and high-level feature maps from various layers in the model, including convolutional base, pooling, and batch-normalization layers. These features are then used to create a one-dimensional feature vector, which is

inputted into fully connected layers for classification. However, CNNs require a substantial amount of annotated data to train effectively and mitigate issues such as underfitting and overfitting. Transfer learning addresses this challenge by leveraging knowledge from pre-trained models, originally trained on large benchmark datasets like ImageNet, and adapting them to new problems with fewer data points, such as classifying lung cancer from CT-scan slices. Fine-tuning involves retraining the top layers of pre-trained CNN architectures on specific tasks, adjusting them to the target domain. In this study, transfer learning is applied to five variants of pre-trained EfficientNet models (EfficientNet B0-B4), fine-tuning them on CT-scan slices of lung cancer for classification. The feature maps extracted from EfficientNet are then utilized in fully connected layers for classification, with further details on optimizing the classification layers provided in subsequent sections.

Classification of fine-tuned EfficientNet. In 2019, the Google AI research team introduced the EfficientNet series, including EfficientNetB0–EfficientNetB7, as a solution to the issue of overparameterization in deep CNN-based architectures. These models outperform many state-of-the-art architectures such as Inception-V3, ResNet50, Inception-ResNetV2, and DenseNet in tasks like image classification from ImageNet, segmentation, and transfer learning. The EfficientNet series achieves this by utilizing fixed sets of scaling coefficients through a uniform compound scaling approach, instead of arbitrarily increasing network width, depth, and resolution. This approach improves network efficiency while maintaining or even enhancing accuracy.

This model focuses on transfer learning using five pre-trained EfficientNet models (EfficientNetB0 to EfficientNetB4) originally trained on the ImageNet dataset. These models are fine-tuned using CT scan slices of lung cancer. The network architecture of the modified EfficientNetB1 is illustrated in Figure 14. The base model is initialized with ImageNet weights and fine-tuned. To reduce dimensionality, a Global Average Pooling (GAP) layer is added on top of the EfficientNet backbone, while the weights in the convolutional base of each block remain fixed. Additionally, a dropout layer with a probability of 0.5 is inserted after the GAP layer to prevent overfitting. The output layer with 1000 units is replaced by an output layer with 3 units and a Softmax activation layer to accommodate the dataset's three class labels. The entire architecture is re-trained using the IQ-OTH/NCCD lung cancer dataset.

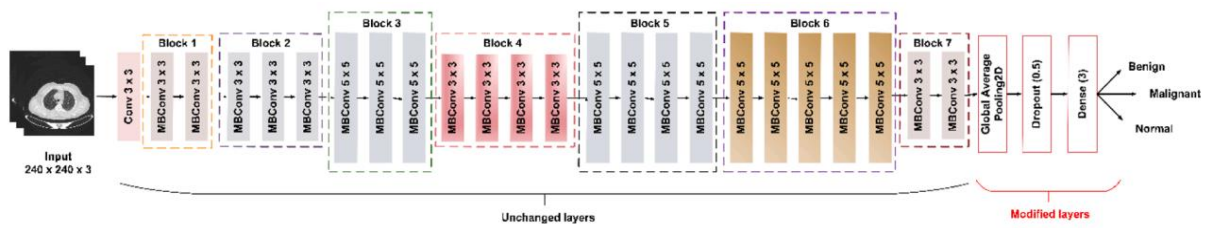


Figure 14 The proposed method of fine-tuning pre-trained EfficientNetB1. The classification layer of the pre-trained EfficientNet is modified by the addition of GAP, dropout, and output layers. The layers in each EfficientNet block are kept unchanged.

With Image pre-processing

The above-mentioned traditional lung cancer prediction techniques fail to achieve higher accuracy this is because of low-quality image which affects the segmentation process. To overcome this issue new optimized image processing and machine learning technique is introduced to predict the lung cancer.

For the considered above dataset images are collected, are examined by applying multilevel brightness-preserving approach which effectively examines each pixel, eliminates the noise, and increase the quality of the lung image. From the noise-removed lung CT image, affected region is segmented by using improved deep neural network that segments region in terms of using layers of network and various features are extracted. Then the effective features are selected with the help of hybrid spiral optimization intelligent-generalized rough set approach, and those features are classified using ensemble classifier. The discussed method increases the lung cancer prediction rate.

CT image preprocessing using multilevel brightness-preserving approach:

The first level of the work is to removing noise from the captured CT lung image because image capturing process consists of several unwanted information, radiation processing information and patient details that occupy the captured X-ray image. The unwanted details reduce the efficiency lung cancer prediction system. Then the noise present in the image is eliminated with the help of multilevel brightness-preserving approach that examines every pixel present in the captured X-ray images effectively. The introduced method effectively examines each image and its pixel for enhancing the quality of image effectively.

Once the image brightness is lower to the neighbouring pixels, the pixel value is replaced by utilizing the mean value of pixel. While doing this process, image is divided or partition into different sub-images, each sub-image is analysed separately and normalizes the image effectively. Considering the input lung CT image I , which is decomposed into two different sub-images according to the image brightness value μ_i , the decomposed images are named as the

foreground image and background image that are denoted as, I_f and I_b . From the initialization, the image is represented as follows.

$$I = I_b \cup I_f \quad (1)$$

In Eq. (1), the image pixels are represented as m and n that are further expressed as follows,

$$I_b(m, n) = \{I((m, n)|I(m, n)) < \mu_I, \forall I(m, n) \in I\} \quad (2)$$

The foreground sub-images are represented as follows,

$$I_f(m, n) = \{I((m, n)|I(m, n)) \geq \mu_I, \forall I(m, n) \in I\} \quad (3)$$

$$p(I_k) = \frac{n^k}{n} \quad (4)$$

In Eq. (4), k having value from 0 to $L - 1$ where L is the number of pixels in image, n^k is total times of level in image I and n is amount of input lung cancer sample images. After that, cumulative density value is computed as follows,

$$C(I) = \sum_{j=0}^k p(I_j) \quad (5)$$

The computed probability density and cumulative density functions are mapped with the histogram of the lung CT image. According to the computation, the histogram of the image is determined as follows,

$$f(I) = I_0 + (I_{L-1} - L_0)C(I) \quad (6)$$

In Eq. (6), $f(I)$ is defined as the transformation function of the input image and $C(I)$ is cumulative function of image. Based on the discussion, the output of the pre - processed image is defined as,

$$Y = f(I) = \{(f(I(m, n))|\forall I(m, n) \in I)\} \quad (7)$$

This above process is applied on both foreground and background divided sub-images effectively. This process also gradually improves the quality of the image, eliminating the noise from the captured CT image. Based on the above discussion, the quality-enhanced image eliminated from the noise is shown in figure 15.










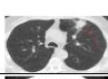
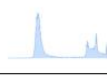
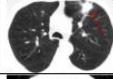

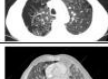

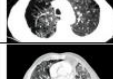




S. no	CT lung image	Noise removed image	Histogram image	Enhanced image
1				
2				
3				
4				
5				

Figure 15 Noise Removed CT Image

The second step is to extract the affected region from noise-removed lung CT image that is commonly known as segmentation. In this work, the segmentation is done by applying improved EfficientNet it utilizes the multiple layers while processing the image. During the semantic segmentation analysis, the neural network follows several steps such as pixel classification because it effectively examines each pixel and predicts whole inputs present in the lung CT image. Which is explained in the below framework clearly figure 16.

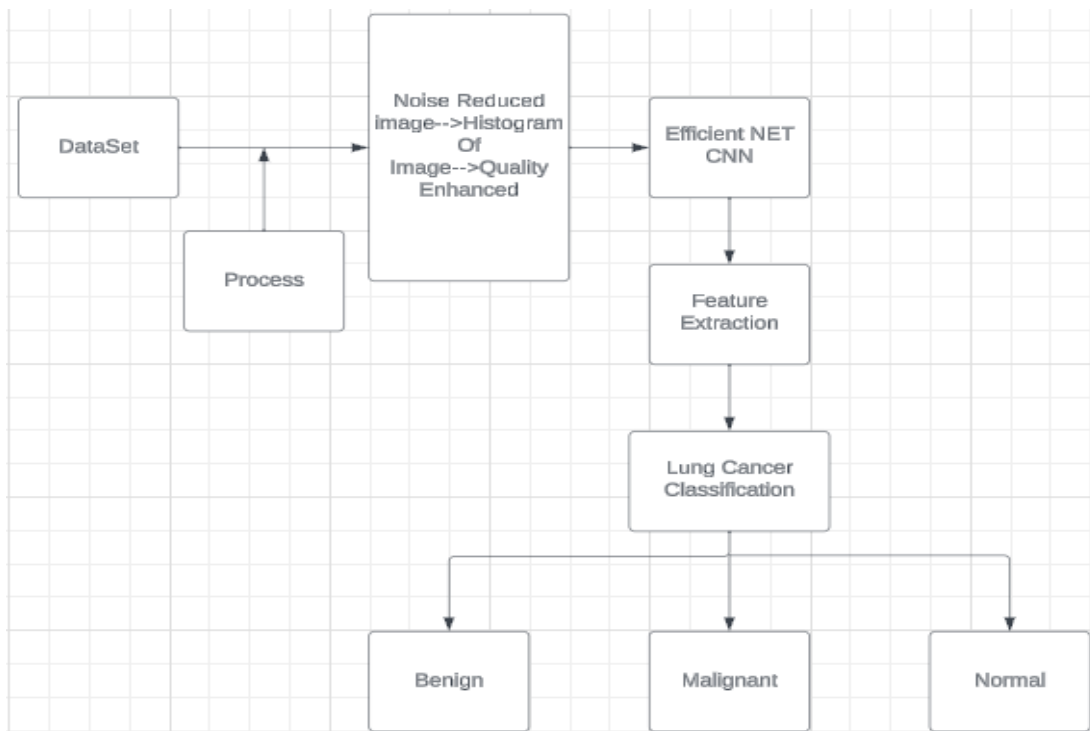


Figure 16 Proposed framework for Lung Tumor Detection with Image pre-processing techniques

The Out-Comes of the model combining both the methods pre-processing and original model is present in the discussion section under the same Lung Tumor Detection.

5.4 Covid-19 Detection

Since the arrival of Covid-19, many researches have been initiated for its accurate prediction across the world. Pneumonia is closely related to covid-19, as several patients died due to chest congestion which is a pneumonic condition. It was challenging for medical experts to differentiate between covid-19 and pneumonia [51].

COVID-19 spreads more swiftly than typical influenza cases, making it challenging to control, despite efforts for early detection. Although the widely used RT-PCR method is commonly employed for COVID-19 diagnosis, it may yield false negative results.

The convolutional neural network (CNN) stands out as a widely embraced deep learning method, increasingly recognized for its capacity to revolutionize various aspects of human existence. This study endeavors to create advanced CNN models incorporating conceptual transfer learning for the detection of COVID-19 using CT scan images. Despite working with limited datasets, these methodologies have shown promising effectiveness in identifying the presence of COVID-19.

The dataset contains 2,481 CT scan images of the lungs sourced from Kaggle. CT scans, a type of X-ray, are utilized for precise diagnosis of internal organs. The dataset is divided into two categories: COVID and non-COVID. CT images of COVID patients are categorized under COVID, while those of healthy individuals are under non-COVID. The COVID class comprises 1,252 CT images, and the non-COVID class includes 1,229 CT images. Along with those SPGC-COVID dataset, which includes CT scans from individuals with both COVID-19 and CAP, as well as scans from uninfected individuals. This dataset provides labels at both the slice-level and patient-level for numerous individuals.

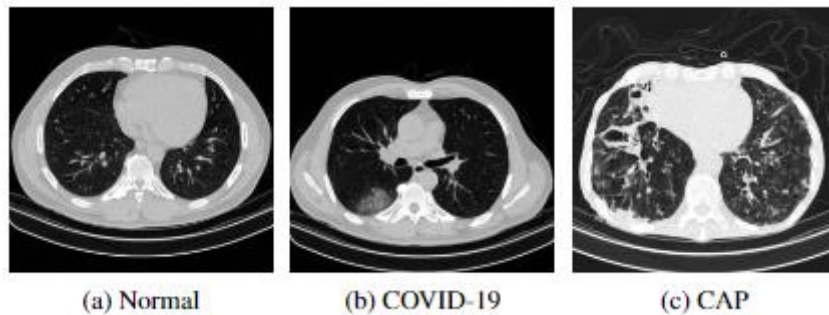


Figure 17 Segmented slice Images from three example CT-Volumes of a) Normal, b) Covid-19, c) CAP patients

Without Image pre-processing

Deep learning, an emerging field, has witnessed numerous successful applications across various domains in recent years. Its use has made solving artificial intelligence problems notably more manageable. Among its applications, deep learning, especially Convolutional Neural Networks (CNN), has gained widespread preference for image classification due to its high accuracy and ability to extract features effectively. CNN has been extensively employed in image analysis since its inception and has shown promise in the medical imaging domain. Various deep CNN architectures, including VGG16, VGG19, InceptionV3, ResNet50, DenseNet121, and Xception, are being explored. This model begins by thoroughly analyzing individual baseline models, where the convolution components remain consistent with standard models proposed in the ImageNet challenge. These models are then applied to chest CT images, with images of varying sizes resized to uniform dimensions before inputting them into the model.

The DenseNet121 design improves upon ResNet by introducing dense connections, where each layer is connected to all subsequent layers. This highly linked architecture ensures that each layer receives network parameters from the layers above it and passes its feature map to the layers below it. This structure allows for feature reuse while minimizing the overall number of parameters. Commonly used DenseNet versions, such as DenseNet121, are employed in the study. The ResNet50 architecture, addresses the vanishing gradient problem in deep neural networks by incorporating shortcut connections. Convolutional and pooling layers are stacked on top of each other, which can lead to performance degradation. However, the use of residual blocks with skip connections effectively mitigates this issue. ResNet50, a variant of ResNet, comprises 50 layers in its architecture.

The InceptionV3 architecture aims to mitigate the issue of significant variations in the positioning of prominent features in images by incorporating multiple kernel types at each level, thereby broadening the network. Inception modules enable the simultaneous operation of numerous kernels. InceptionV3 is an advancement of InceptionV2, addressing concerns regarding representational bottleneck. It introduces auxiliary classifiers with kernel factorization and batch normalization, enhancing the network's performance. The Xception network utilizes depth-wise separable convolution layers and consists of 36 convolution layers grouped into 14 modules. Each layer is connected by continuous residual links, with the first and last layers excluded. The model focuses on mapping spatial and cross-channel correlations, which can be independently handled in CNN feature maps. The fundamental design of Inception is retained in Xception, emphasizing spatial correlation

transformation within each output channel to capture cross-channel correlations.

The VGG Net, devised by Simonyan and Zisserman from the Visual Geometry Group at the University of Oxford, is a well-known Deep CNN model. It achieved top rankings in the ILSVRC 2014 image classification tasks, with first and second places. To enhance accuracy in Computer Vision tasks, the models incorporate increased layers of CNN with a few kernels. VGG architecture variations have found wide application in extracting deep image characteristics, especially in the medical domain. VGG 16 comprises sixteen layers, while VGG19 consists of nineteen layers.

After taking the images from dataset, approximately 20% of the dataset was reserved for testing purposes, while the remaining 80% was used for training. Convolutional Neural Networks (CNNs) represent a significant advancement in image recognition, particularly for visual imagery analysis and classification tasks. They are commonly implemented using Keras, a Python library with a TensorFlow backend, which is a deep learning framework. Due to its extensive dataset (1.2 million images), ImageNet has been frequently utilized to develop various CNN architectures. To extend the model's applicability beyond the ImageNet dataset, transfer learning is employed, leveraging pre-trained models. Fine-tuning techniques are also utilized to adapt pre-trained models to specific tasks or domains.

Our model was constructed using various pre-trained models like Densenet121, InceptionV3, Xception, ResNet50, VGG16, and VGG19. For example, the VGG19 convolutional neural network (CNN) with a 2D Max Pooling output layer of shape (7, 7, 512) and no parameters can be imported using the Keras applications class. Since we only have two classification classes, Covid19 and Normal (No-Covid19), we applied transfer learning to adapt the output layer into a binary classifier. The outputs from the preceding layer were retrieved and passed to a Flatten layer, resulting in a flattened array with dimensions of $7 \times 7 \times 512 = 25088$. As rectangular or cubic shapes cannot be directly used as inputs, this flattening process is typically performed towards the end of the CNN. To prevent overfitting and improve generalization, we added a Dropout layer as a regularization strategy, with a dropout rate of 0.5, which is a commonly used value. This decision is based on the observation that large neural networks on small datasets often overfit the training data, leading to a decrease in validation accuracy.

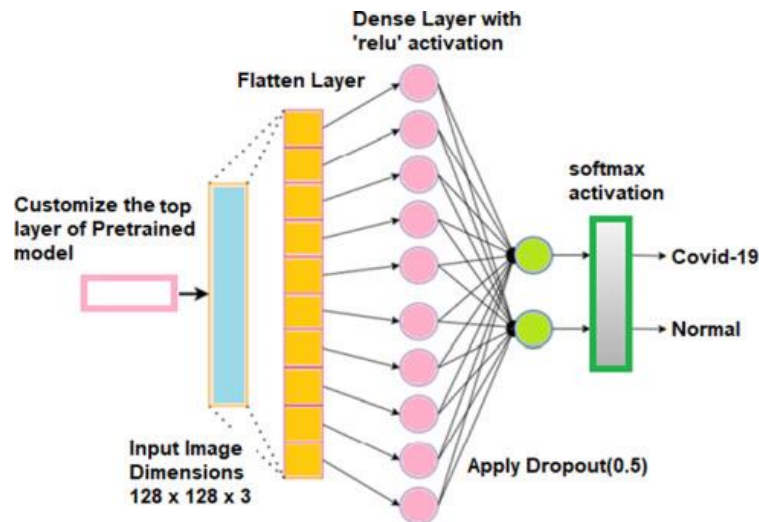


Figure 18 Model of a custom structure which is used in main framework

With Image pre-processing

We aim to focus on early detection and classification of lung diseases from the raw X-ray images for appropriate treatment using the semi-automated approach for robust feature extraction and deep learning with minimum computation overhead. The requirement of robust and reliable detection and classification of lung diseases (like Covid-19 and viral/bacterial pneumonia) motivates the Image Pre-processing technique.

The input raw X-ray image is pre-processed by applying the median filtering and histogram equalization techniques. To extract the region of interest (ROI) from the enhanced image, a dynamic region growing technique for the image of different modalities and dimensions has been proposed. The sets of feature vectors have been built from the ROI images using visual, texture, intensity, and invariant moment features. • The robust set of features such as Histogram of Oriented Gradient (HOG), texture, invariant moment and ROI intensity features have been extracted from the segmented image. After that, we apply the feature scaling (normalization) method to enhance detection accuracy. The normalized features can be used for disease severity analysis by medical practitioners effectively.

X-ray image quality enhancement:

As the quality of input chest X-ray scans is low and contains noise, it may lead to suppressions of lung regions affected by congestions or fluids. Most recent techniques directly applied the deep learning models without quality enhancement; however, such methods may not be reliable for a longer time. The input chest X-ray image I has been pre-processed in the proposed model F-RRN-LSTM using adaptive intensity values adjustment, median filtering, and histogram equalization. The first operation focused on the adjusting the

image intensity values for low contrast X-ray images. This technique mainly used to enhance the contrast as:

$$I^1 = imadjust(I) \quad (1)$$

where, I^1 is outcome of contrast enhancement step using function *imadjust*.

The median filtering has been applied further to remove the noise in the contrast-enhanced image. Adjusting the image intensity values leads to noise and X-ray scan introduces the noise in the image. Median filtering shows the effective enhancement in comparison to the adaptive bilateral filtering, average filtering, and wiener filtering for X-ray datasets. Median filtering is a lightweight technique and commonly used in many image processing applications as it is more effective while considering the constraints of noise reduction and edge preservation. We used 2D median filtering in the proposed model that works by moving via the image pixel by pixel, replacing every value with the neighbouring pixel's median value. The neighbour's pattern is decided by the size of the window. The window size of a 3×3 neighbourhood has been used in this work. The 2D median filter is applied on I^1 as:

$$I^2(i, j) = median\{I^1(i, j)\}(i, j) \in w \quad (2)$$

where, I^2 is outcome of the median filtering and w is the size of window.

The outcome of the image enhancement function in the proposed model is demonstrated on two sample chest X-ray images as shown in Fig. The poor-quality X-ray images have been enhanced effectively by persevering the edges and removing the noisy portions. We have also tested the other filtering approaches (adaptive bilateral filtering, average filter, and wiener filter) and contrast enhancement techniques (histogram equalization and contrast limited adaptive histogram equalization), but these suffered from loss of edges information, increased background noise contrast, and lung regions data loss. The outcome shows the raw X-ray images improved with optimal contrast and image quality. This enhancement helps to accurately detect the lung regions during the ROI extraction process.

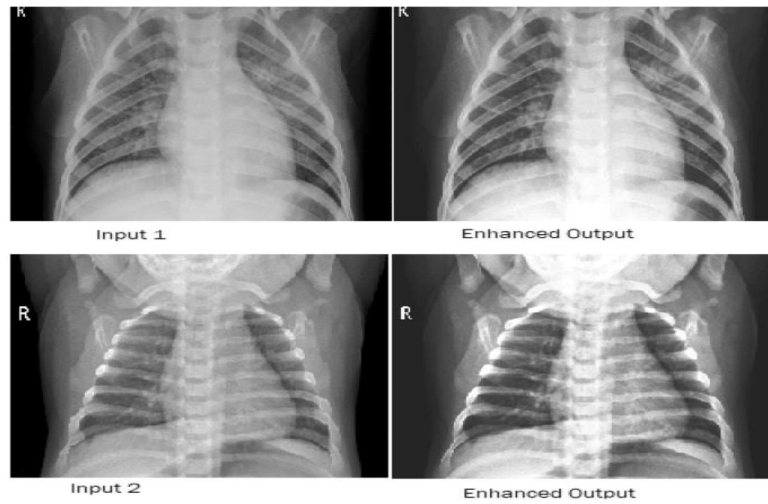


Figure 19 Outcomes of Chest X-Ray Image Pre-Processing

The above Image pre-processing technique is combined with the above mentioned covid-19 model and is designed below figure 20.

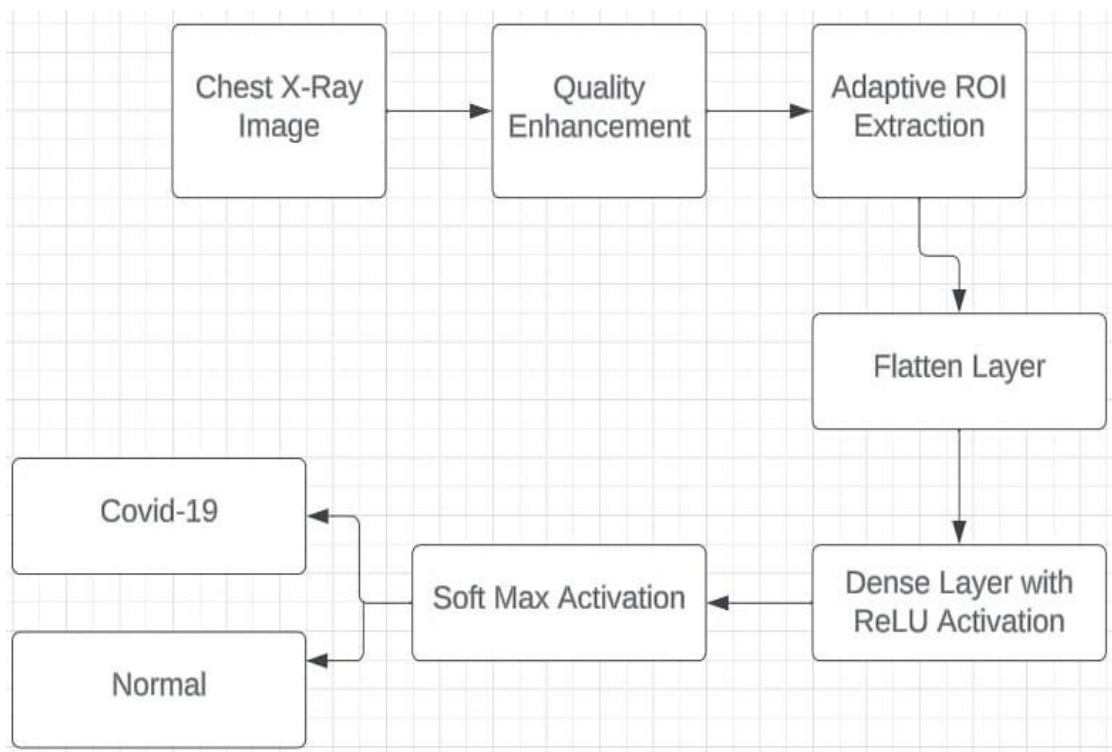


Figure 20 proposed framework for Covid-19 detection containing Image Pre-Processing techniques

The Output of the model combining both the methods pre-processing and original model is present in the discussion section under the same Lung Tumor Detection.

5.5 Medical Report Analysis

Efficiently obtaining insights into infectious diseases is crucial for population health research. However, analyzing extensive health data faces obstacles due to the absence of protocols. A study aimed to use natural language processing (NLP) to extract essential details, including clinical factors and social determinants of health, from unstructured text. The proposed framework involves creating a database, NLP modules for identifying clinical and non-clinical information, and an evaluation protocol. Combining NLP and machine learning techniques, the study extracted information from documents like clinician notes, radiology reports, and lab reports to generate structured variables for real-world data analysis. Using a large electronic health record (EHR) database, models were chosen based on performance. Variables curated through ML extraction are determined solely by ML models, without confirmation by abstraction, identifying crucial information from visit notes and documents without predicting future events or filling in missing information.

A novel technique combining natural language processing (NLP) and machine learning (ML) was developed to extract clinically relevant data from unstructured electronic health record (EHR) documents. This approach showed exceptional performance in extracting key variables, including details such as the date of initial cancer diagnosis, disease stage, histology, smoking status, surgery dates, biomarker test results with corresponding dates, and dates of oral treatments. Compared to manual curation, which is time-consuming and resource-intensive, this method offers efficiency and scalability, enabling larger patient cohorts for research. Additionally, it addresses challenges associated with outdated information and limited sample sizes, which can hinder critical research efforts.

This model outlines the process for using natural language processing (NLP) and machine learning (ML) at Flatiron Health to extract information from unstructured documents within oncology care electronic health records (EHRs). It emphasizes the distinction between "abstraction" tasks performed by humans and "extraction" tasks handled by models.

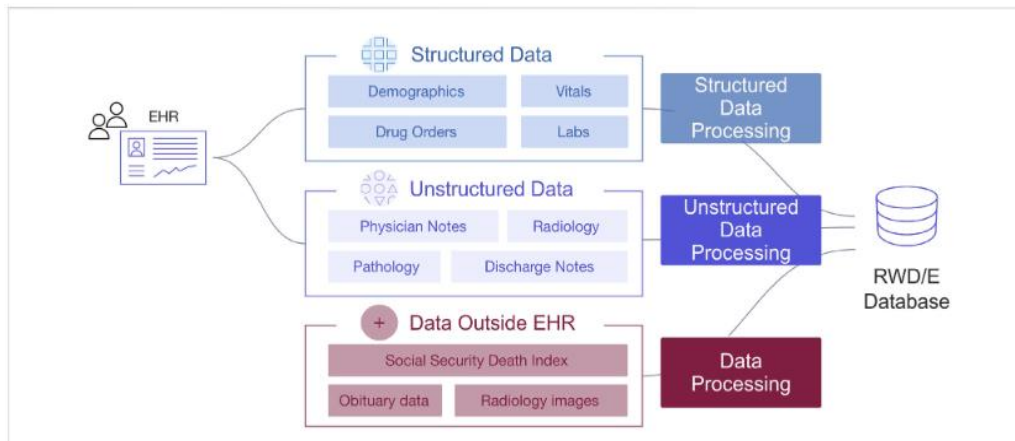


Figure 21 Overview of data variables defined by structured and unstructured information in EHR.

Clinical experts compile a list of relevant clinical terms and phrases to identify variables of interest from electronic health record (EHR) documents. Natural language processing (NLP) techniques locate sentences containing matches to these terms, with optical character recognition (OCR) systems extracting text from various document formats. Machine learning (ML) models analyze the presence of clinical concepts and contextual information in surrounding snippets, along with associated dates. The sentences are transformed into a mathematical representation for interpretation, resulting in a comprehensive set of features capturing document structure, chronology, and relevant clinical terms or phrases.

The training set consists of labeled data points for optimizing the model's parameters. It undergoes iterative processes where the model is trained with examples, its outputs are compared to labels, and parameters are adjusted to minimize errors. The objective is to teach the model to predict answers akin to human abstractors when analyzing clinical text. The validation set, not encountered during training, evaluates the model's learned associations and estimates its performance on new, unlabeled examples. Evaluation metrics like precision, recall, and F1 score, along with error analysis, guide decisions on search terms, preprocessing, and model architectures. In classification tasks, models output scores for each class based on electronic health records, with the highest score often determining the assigned class. Empirically chosen thresholds may optimize performance, considering class balance. Alternatively, abstraction can resolve uncertainty when no class receives a sufficiently high score. Various ML models like LSTM, GRU, and BERT are explored to extract variable information from electronic health records, leveraging contextual understanding for nuanced text analysis. The model framework is built below figure 22.

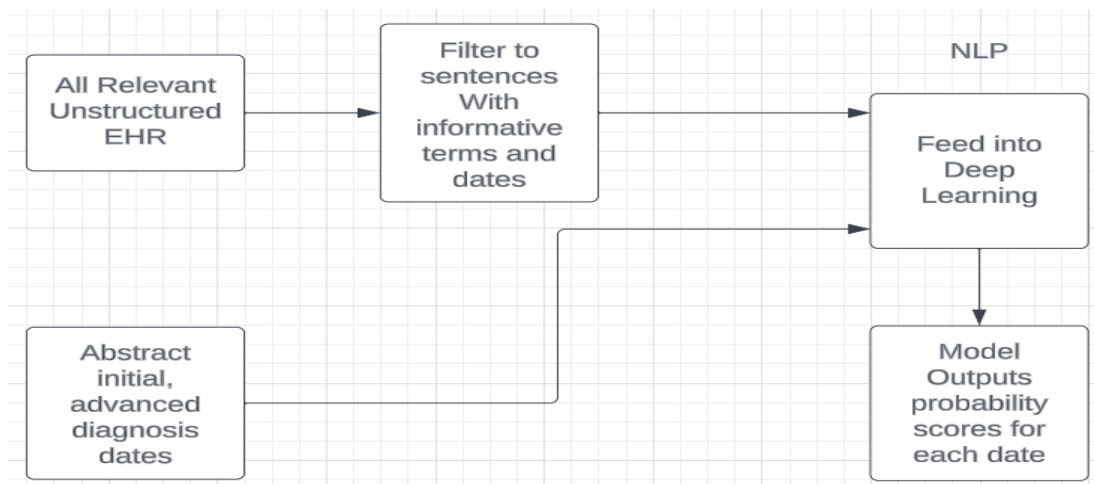


Figure 22 Proposed framework for Medical Report Analysis containing image pre-processing techniques

The results of this model are still under development so, there will be no discussion for the methodology of Medical Report Analysis.

6. Methodology of Gen-AI in Medical Field

GEN-AI, or Generative Artificial Intelligence, has transformative potential in the medical field. It encompasses AI techniques like generative adversarial networks (GANs) and deep learning to create synthetic data, images, or text that closely resemble real-world examples. In medicine, GEN-AI is applied for tasks such as medical image synthesis, data augmentation, and generating realistic patient data for training models. It helps address challenges like limited data availability, privacy concerns, and data imbalance. Additionally, GEN-AI facilitates the creation of diverse datasets, improving the robustness and generalization of AI models in healthcare applications.

6.1 Gen-AI in Drug discovery and development

Target Identification and Validation: Identify a specific molecular target associated with a disease. Analyzing biological networks to identify key nodes and pathways associated with diseases. Incorporating genomics, proteomics, and other omics data to identify potential therapeutic targets.

High Throughput Screening (HTS): Use automated methods to quickly test the biological activity of thousands of compounds. Utilizing novel assay formats, such as biosensors and microfluidics, for more efficient screening. Implementing machine learning algorithms to analyze large datasets generated from HTS.

Lead Identification and Optimization: Select and optimize promising compounds for further development. Using machine learning models to

predict and optimize chemical properties for better drug-like characteristics. Investigating how changes in the chemical structure of a compound impact its biological activity.

Preclinical Testing: Assess the safety and efficacy of the selected compounds in laboratory and animal studies. Mimicking human organ microenvironments in microfluidic devices for more predictive preclinical studies. Utilizing advanced imaging technologies to monitor drug distribution and effects in live animals.

Clinical Trials: Conduct rigorous human trials to evaluate safety, dosage, and efficacy. Modifying trial parameters based on interim results to enhance efficiency and flexibility. Incorporating patient feedback and experiences for more patient-friendly and inclusive trials.

Fragment-Based Drug Design (FBDD): Utilizes small fragments of molecules to identify lead compounds. Expanding and optimizing initial fragment hits to improve binding affinity and selectivity. Employing high-resolution structural techniques to visualize fragment interactions with target proteins.

In Silico Modeling: Computer simulations and modeling to predict drug behavior in biological systems. Combining quantum mechanics for accurate electronic structure calculations with molecular mechanics for larger-scale molecular dynamics simulations.

Pharmacogenomics: Considers how an individual's genetic makeup influences their response to drugs. Tailoring drug treatments based on individual genetic variations for more effective and safer outcomes.

High-Content Screening (HCS): Combining automated microscopy with quantitative image analysis for detailed cellular studies. Implementing advanced algorithms to extract meaningful information from complex cellular images.

Virtual Screening: Employing computational methods to prioritize compounds for experimental testing. Comparing methods that focus on known ligands or utilize structural information of target proteins.

Cryo-Electron Microscopy (Cryo-EM): Visualizing biomolecules at near-atomic resolution, aiding in structure-based drug design. Differentiating between techniques used for studying isolated particles and cellular structures.

6.2 Gen-AI in Diagnosis

Medical Imaging: Techniques like X-rays, CT scans, MRI, and ultrasound for visualizing internal structures. In addition to structural imaging, incorporating functional modalities like PET (Positron Emission Tomography)

and fMRI (Functional Magnetic Resonance Imaging) to assess physiological processes.

Laboratory Tests: Analyzing blood, urine, and other samples for markers of diseases. Utilizing portable devices for rapid on-site analysis, allowing for quicker treatment decisions.

Genetic Testing: Examining an individual's genes for genetic disorders or susceptibility to certain diseases. Enabling rapid and cost-effective sequencing of entire genomes for comprehensive genetic analysis.

Clinical Examination: Physical examination and patient history to form a diagnosis. Integrating electronic health records to provide a comprehensive view of a patient's medical history.

Liquid Biopsy: Detecting and analyzing circulating tumor cells or cell-free DNA for cancer diagnosis. Identifying specific genetic mutations in cell-free DNA to guide targeted cancer therapies.

Telemedicine and Remote Monitoring: Using technology to provide healthcare services remotely. Utilizing wearables for continuous monitoring of vital signs and health parameters.

Metabolomics and Proteomics: Profiling metabolites and proteins for comprehensive disease diagnostics. Profiling small molecules to understand metabolic pathways and identify disease-specific markers.

Digital Pathology: Digitizing and analyzing pathology slides for more accurate and efficient diagnosis. Digitizing entire pathology slides for remote viewing and collaboration among pathologists.

6.3 Gen-AI on Remedies

Pharmacotherapy: Using drugs and medications to treat diseases. Tailoring drug treatments based on individual patient characteristics, including genetics and biomarkers.

Surgery: Invasive procedures to remove, repair, or replace affected tissues. Techniques like laparoscopy and robotic-assisted surgery for reduced invasiveness and faster recovery.

Physical Therapy: Rehabilitation and exercises to restore physical function. Tailoring exercise programs to specific conditions for optimal recovery.

Alternative and Complementary Medicine: Non-conventional approaches like acupuncture, herbal medicine, etc. Understanding the principles and mechanisms of acupuncture in pain management and other health conditions.

Immunotherapy: Stimulating the immune system to fight diseases like cancer. Understanding the role of drugs that enhance the immune system's ability to recognize and attack cancer cells.

Regenerative Medicine: Using stem cells and tissue engineering to repair or replace damaged tissues. Exploring the potential of stem cells for tissue repair and regeneration.

Nanomedicine: Designing and delivering drugs at the nanoscale for targeted therapy. Nanoparticles, liposomes, and other nanocarriers for targeted and controlled drug delivery.

CRISPR/Cas9 Gene Editing: Precision gene editing for potential correction of genetic disorders. Exploring the potential of CRISPR/Cas9 for correcting mutations responsible for genetic diseases.

6.4 Gen-AI on Medication (Drug Prediction)

Computational Drug Discovery: Utilizing AI and machine learning algorithms to predict drug-target interactions. Using algorithms to generate entirely new molecular structures with desired properties for specific targets.

Chemo informatics: Analyzing chemical data to understand the structure-activity relationship of drugs. Extracting valuable information from large chemical databases to identify trends and patterns.

Quantitative Structure-Activity Relationship (QSAR): Modeling the relationship between chemical structure and biological activity.

Deep Learning in Drug Discovery: Harnessing neural networks for more intricate and accurate predictions. Designing deep learning models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for drug discovery tasks.

Pharmacophore Modeling: Identifying common features essential for drug-receptor interactions. Ensuring the reliability of pharmacophore models through rigorous validation techniques.

Adaptive Clinical Trial Designs: Dynamic approaches that allow for adjustments based on interim data analysis.

Network Pharmacology: Studying the interactions between drugs and biological networks for a holistic understanding.

6.5 Gen-AI in Virtual Health Care Assistants (Chatbots)

Natural Language Processing (NLP): Enabling machines to understand and generate human-like language. Analyzing the structure and meaning of sentences to comprehend user queries more accurately.

Machine Learning Algorithms: Training chatbots to learn and improve their responses over time. Allowing chatbots to learn from user interactions and feedback to continually enhance their performance.

Integration with Electronic Health Records (EHR): Accessing and updating patient information. Providing chatbots with access to the latest patient information to offer contextually relevant responses.

Sentiment Analysis: Assessing the emotional tone in patient communication for more empathetic responses. Detecting signs of distress or urgency in the language to prioritize and escalate cases as needed.

Multimodal Conversational Agents: Integrating text, voice, and visual cues for a more natural interaction. Analyzing visual information for a comprehensive understanding of user queries, such as skin condition assessments or medication identification.

Transfer Learning in NLP: Leveraging pre-trained models for improved performance in healthcare-specific language. Adapting general language models to the specialized vocabulary and nuances of the healthcare industry.

Personal Health Record Integration: Accessing and updating patient information seamlessly during interactions. Ensuring that systems adhere to interoperability standards like HL7 to facilitate the exchange of health information.

7. Discussion on Image Pre-Processing and Medical Diagnosis

The outputs, observations, the evaluations that have been performed on our model have been presented in this section.

The below table 1 depicts the state-of-the art models which have infused many image enhancement techniques and classification models which have been used in the researches that have been presented in the present field.

The evaluation of the proposed methods has been done by their accuracy, precision, recall, and f1 score.

Precision

Precision is a metric that measures how robust enough the model recognizes the positive samples and is specific to the expected outcome. The better value of precision shows, the more precise the positive sample prediction. Eq. (1) represents the precision formula for calculation.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Recall

Improved accuracy in predicting the target instance and decreased likelihood of missing a bad instance are both the result of a higher recall rate. The method to find recall is expressed in Eq. (2).

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

Accuracy

Accuracy is defined as the proportion of correct predictions made relative to the total number of samples. Eq. (3) is used for calculating the Accuracy of the various models.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

F1-Measure

The F1 score is offered as a composite metric to mitigate the negative effects of both precision and recall when evaluating classifiers. F1-score is calculated using Eq. (4).

$$F1 - score = 2 \left(\frac{Precision * Recall}{Precision + Recall} \right) \quad (4)$$

Table 1 Evaluation of the start-of-the art models which encapsulate image enhancement techniques in their original model

Task	Motion Correction	Instrumental artefacts correction	Sequence artefacts correction	Multi-site normalization	Noise Reduction	Super-resolution
Scopes	Rigid motion, non-rigid motion	Gibbs ringing, B0 inhomogeneity	EPI ghosting and distortion	Multi-site normalization	Anatomical	Anatomical, DWI, spectroscopic
Encoder Types	2D Conv	2D Conv, 3D Conv	2D Conv, 3D Conv	2D Conv, 3D Conv	2D Conv, 3D Conv	2D Conv, 3D Conv, MLP
Model Types	CNN, Inception, ResNet, FCN, VAE, U-Net	CNN, U-Net, ResNet50	AE, U-Net	CNN, U-Net, CNN with attention, VAE, ResNet	CNN, AE	CNN, U-Net, AE
Training types	Fully supervised, adversarial	Fully supervised, adversarial	Fully supervised, adversarial	Fully supervised, adversarial	Fully supervised, adversarial	Fully supervised, adversarial
Loss	MAE, MSE, adversarial loss	MAE, MSE, adversarial loss	MAE, MSE, SSIM loss, gradient loss	MAE, perceptual loss	MAE, MSE, perceptual loss	MAE, MSE, adversarial loss, SSIM loss, triplet loss

Input Types	2D image, 2D patch, 3D patches	2D image	2D image, 2D patch	2D image	2D image, 2D patch	2D image, 2D patch, 3D patch
Inputs	Images/patches with simulated motion artefacts	Images/patches with simulated Gibbs artefacts or simulated phase errors	Images/patches with ghosting and distortion	Images acquired on different scanners or with different parameters	Images/patches with simulated Gaussian/Poisson/Rician noise	Low-resolution images
Outputs	Motion-free images/patches	Artefact-free images/patches	Artefact-free images/patches	Labels for the task of interest	Noise-free image/patch	High-resolution images
Training Data Source	Motion artefacts are simulated by adding phase error in k-space	Gibbs ringing artefacts are simulated with cropped kspace, created bias fields for simulating B0 inhomogeneity	Artefact-free images are generated by post-processing steps	Multi-site data acquisitions	Simulation by adding noises to the clean image to generate noisy training samples	Simulation by undersampling k-space to create low-resolution training samples
Anatomical Regions	Brain, liver, abdomen, pelvis	Brain, knee, respiration	Brain	Brain	Brain, knee	Brain, prostate, knee, fetal, cardiac, torso
Metrics	SSIM, RMSE, MMI	SSIM, PSNR, RMSE, HFEN, GSR	GSR, MSE, Mutual Information	MAE, DICE,	SSIM, PSNR, IFC	SSIM, PSNR, RMSE, DICE
Examples	MocoNet [73]	DeepResp [74]	S-Net [75]	DeepHarmony [76]	RED-WGAN [77]	SR-q-DL [78]

As seen from the table above there are so many models which are incorporating image enhancement techniques and their outcomes were astonishing, which motivated us to perform image pre-processing techniques on the dataset we are considering.

7.1 Discussion on Brain Tumor Detection

The method that we have proposed by classifying it as two stages which can be clearly understood with the framework built and displayed in the methodology section. The model we proposed is then compared with the existing models that have been performed in the brain tumor detection field as shown in table 2.

Model	Methodology Used	Accuracy
[29] CNN	22-layered CNN structure for brain tumor classification using T1-weighted contrast-enhanced MRI images	96.56%
[79] 3D CNN	Deep multi-scale 3D CNN model for grading brain tumors using volumetric MRI images	96.49%
[80] CNN	CAD system utilizing an 18-layered CNN for brain tumor classification	94.74%
[81] CNN	CNN for direct prediction of tumor grade from imaging data	89.5%-92.98%

Model	Methodology Used	Accuracy
[82] ResNet-34	Pre-trained ResNet-34 CNN model for brain tumor identification from MRI images (Less Data is validated)	100%
[83] CNN	CNN for diagnosing prevalent brain tumor types using provided MRI images	Training: 98.51%, Validation: 84.19%
[84] CNN + SVM	CNN-SVM system for automatic brain tumor diagnosis and localization	CNN (AlexNet): 99.55%
[85] CNN	CNNs (AlexNet, GoogLeNet, VGGNet) for distinguishing between brain tumor types using MRI images	VGG16: 98.69%
Ours	Without Image Pre-processing: CNN model Classification	Classification -1: 99.33% Classification -2: 92.66% Classification -3: 98.14%
	With Image Pre-Processing: Image Enhancement + CNN model	Classification -1: 99.70% Classification -2: 95.43% Classification -3: 99.25%

Table 2 Depicts the comparsion of our model with the existing models out there

As seen from the above table our model without image pre-processing has achieved 99.33% accuracy using the first CNN model. The second CNN model can classify the brain tumor into five brain tumor types as normal, glioma, meningioma, pituitary and metastatic with an accuracy of 92.66%. The third CNN model can classify the brain tumors into three grades as Grade II, Grade III and Grade IV with an accuracy of 98.14%.

After implementing image pre-processing techniques, we have achieved 99.70% accuracy using the first CNN model. The second CNN model can classify the brain tumor into five brain tumor types as normal, glioma, meningioma, pituitary and metastatic with an accuracy of 95.43%. The third CNN model can classify the brain tumors into three grades as Grade II, Grade III and Grade IV with an accuracy of 99.25%. Which clearly indicates the performing image pre-processing have increased the accuracy of the over all model which was classified into three different CNN models.

The output is displayed below figure 23 where the model has evaluated the dataset and then classified that into different types with color referencing.

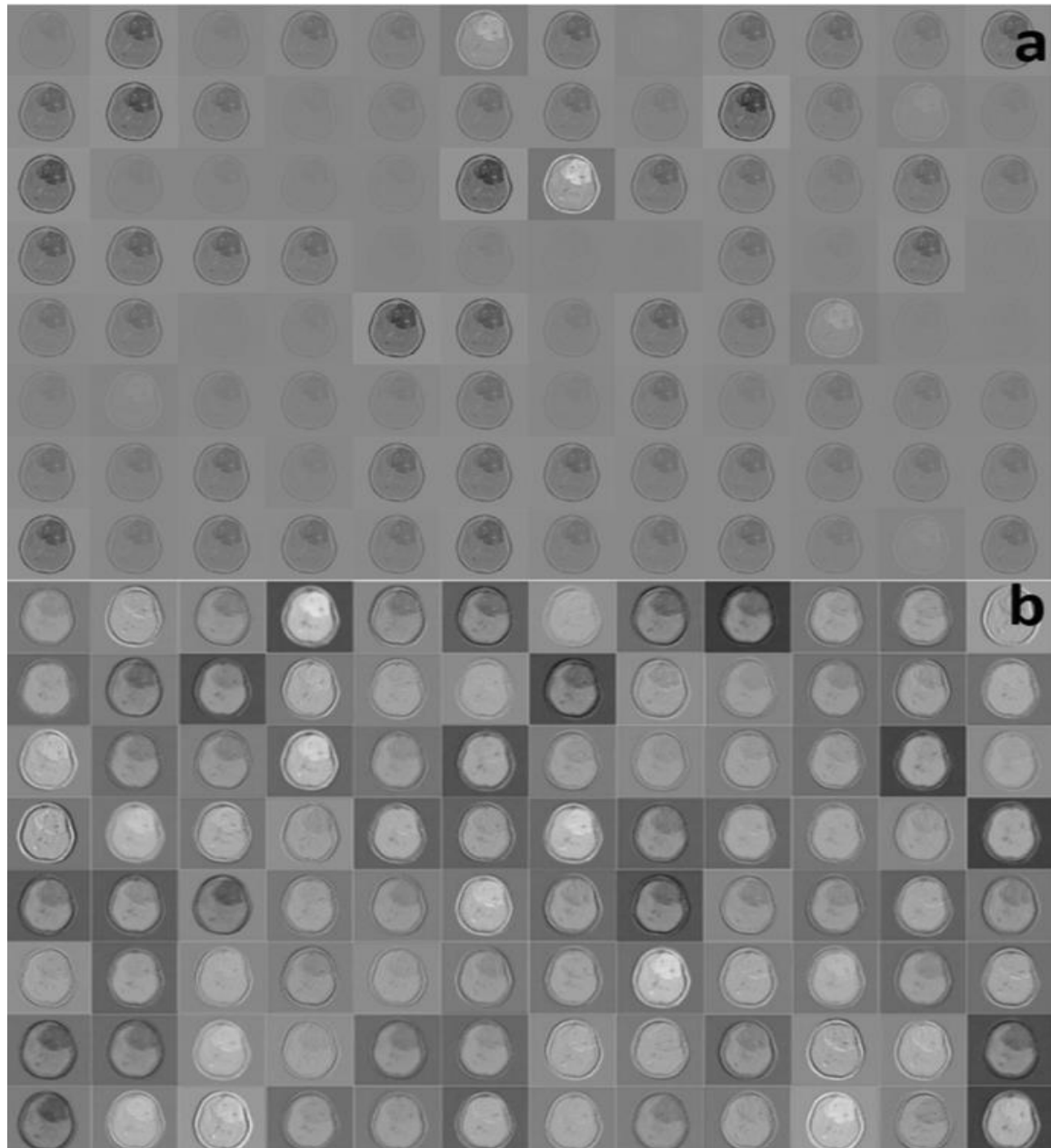


Figure 23 First a and second b convolutional layer activations for Classification-1 task. Each image in the grid is the output of each channel. White regions show strong positive activations, whereas Gray sections show not-strongly activated channels

Now, these images can be generated by any model, but to differentiate from this the proposed framework does all the work from image pre-processing to image classification and generates report to the user in the interface that we have designed which is shown in the below figure 24. Making it easier to understand and doesn't require to consult anyone initially.

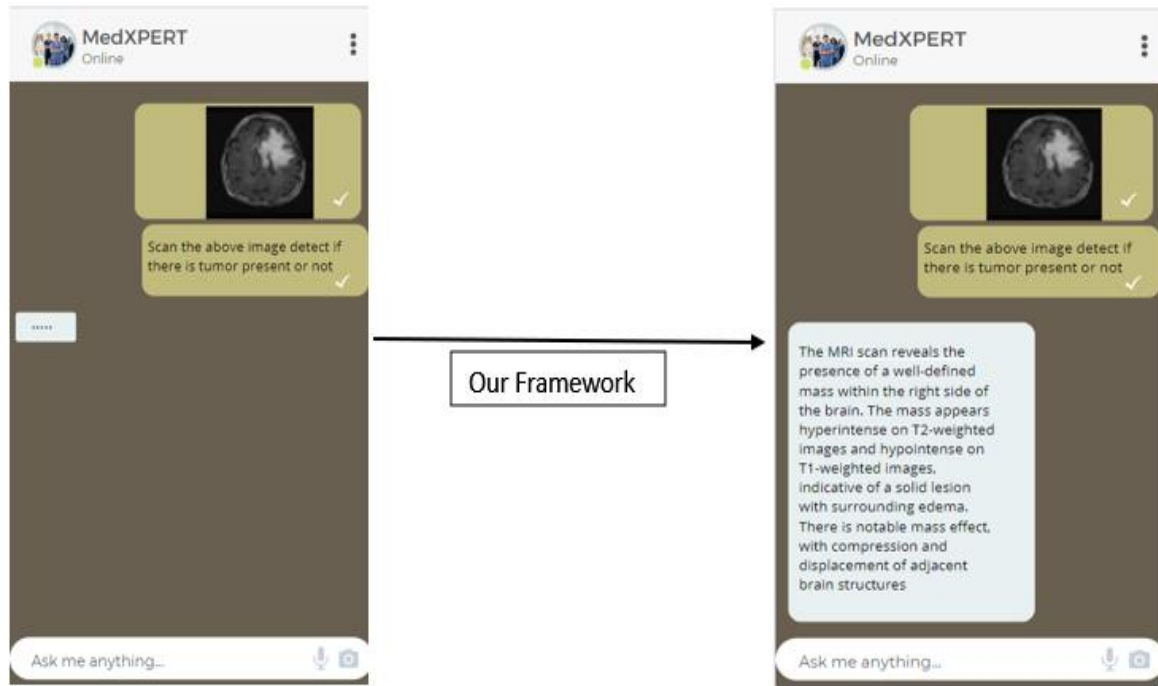


Figure 24 Depicting the proposed interface while working on brain tumor classification and generating result

As seen from the above image our framework is built on a chat interface which takes input from user analyses it and answers accordingly.

7.2 Discussion on Alzheimer Detection

The method that we have proposed by classifying it as two stages which can be clearly understood with the framework built and displayed in the methodology section. The model we proposed is then compared with the existing models that have been performed in the alzheimer detection field as shown in table 3.

Our model undergoes extensive evaluation against DenseNet169, VGG19, and InceptionResNet V2 using metrics such as precision, recall, F1-score, Area Under the Curve (AUC), and loss. The evaluation results for ADD-Net are as follows: 98.63% accuracy, 99.76% AUC, 98.61% F1-score, 98.63% precision, 98.58% recall, and 0.0549% loss. This is achieved without image pre-processing.

After including the image pre-processing methods, the accuracy of the output just increased by a factor of 99.42%, which means a significant development for the main model we used. This can be understood by the framework that has been displayed in the methodology section.

The below table really goes through different models that have been used in the history of Alzheimer disease detection. But, from all those models our model has displayed a higher accuracy in the detection of Alzheimer disease beating all the state-of-the existing models available out there.

Model	Methodology	Used Image Enhancement Techniques (Yes/No)	Accuracy
[35] RF + CNN	Random Forest feature selection combined with CNN classification	No	38.8%
[86] CNN	CNN with structural MRI and DTI focused on hippocampus ROI	No	96.7%
[87] 3D CNN	3D CNN for AD detection with PET scans	Yes	88.76%
[88] RF + CNN	Novel CNN architecture for Alzheimer's severity assessment	Yes	96.7%
[89] Various DL Models	Various DL architectures including 2D/3D CNN and RNN	Yes	96.88%
[90] Transfer Learning CNN	Transfer learning CNN with data augmentation	Yes	99.92%
[91] DenseNet	Classification and segmentation of AD data using DenseNet	No	88.9%
[92] VGG-16, MobileNet	Binary classification of healthy vs. AD brain MR images	No	VGG-16: 93%, MobileNet: 98%
[93] Deep Belief Networks + SVM	Classification of stable vs. progressive MCI data	No	86.6%
[49] SVM	Classification of AD using 2D and 3D brain images	Yes	92.86%
Ours	Without Image Pre-Processing	No	98.63%
	With Image Pre-Processing	Yes	99.42%

Table 3 Comparing our model with state-of-the-art models

The above table just shows the comparison but, whereas the below figure 25. shows the output that has been generated by our model.



Figure 25 Generalization of the class activation map to locate the discriminative region through Grad-CAM.

Now, evaluating this data by a normal person is an impossible task, he/she must refer a practitioner to get it evaluated which is time consuming, and here's where our model comes into play which has state-of-the-art framework to analyse and predict the output from the scans uploaded by user. The whole working is shown in the figure 26 below.

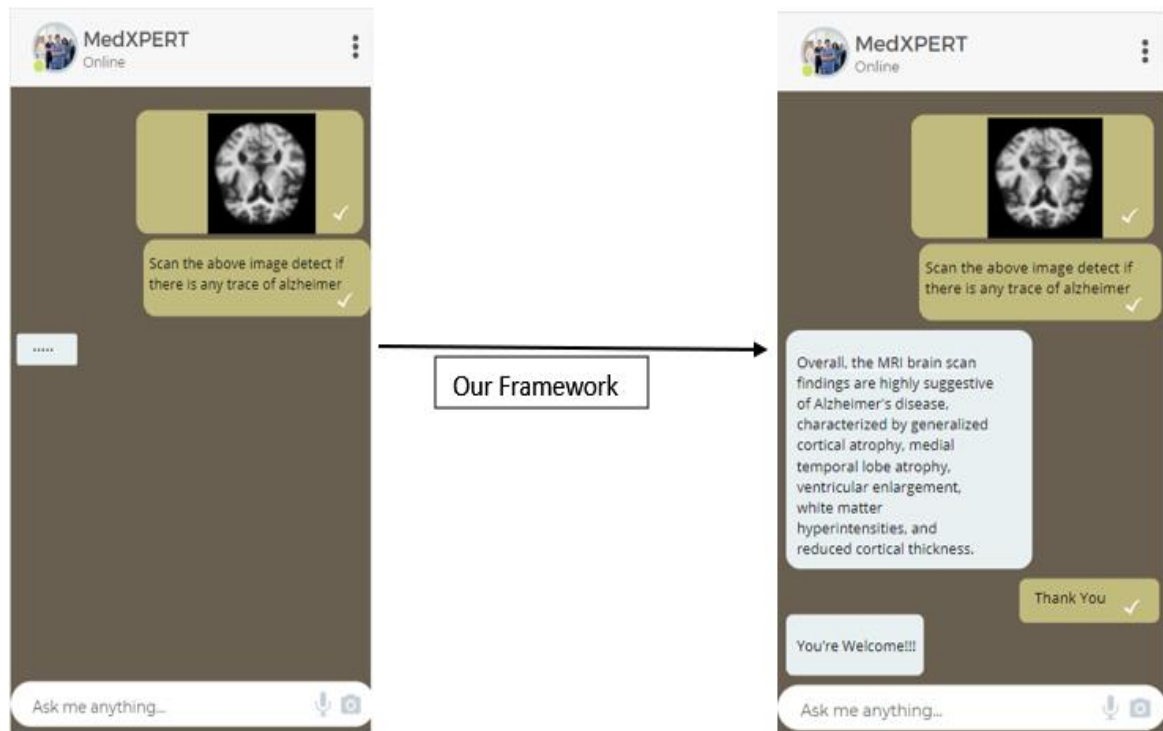


Figure 26 Depicting the proposed interface while working on Alzheimer Disease classification and generating result

As seen from the above image our framework is a chat interface taking input from user and analysing using the framework and responding accordingly.

7.3 Discussion on Lung Tumor Detection

The method that we have proposed by classifying it as two stages which can be clearly understood with the framework built and displayed in the methodology section. The model we proposed is then compared with the existing models that have been performed in the Lung Tumor detection field as shown in table 4.

The evaluation of our model involves the use of five different variants of EfficientNet, namely, B0 to B4. The research involves conducting experiments on the standard dataset "IQ-OTH/NCCD," which includes lung cancer patients categorized as benign, malignant, or normal depending on whether lung cancer is present or absent. To address the class imbalance problem, various data augmentation techniques were employed to mitigate biases. Our model achieved an accuracy of 99.10% and ROC scores ranging from 0.97 to 0.99 on the test dataset.

Model	Methodology Used	Accuracy
[94] VGG16	Transfer learning with data augmentation on IQ-OTH/NCCD lung cancer dataset, retrained on CT-scan images	98.83%
[94] VGG19	Transfer learning with data augmentation on IQ-OTH/NCCD lung cancer dataset, retrained on CT-scan images	98.05%
[94] Xception	Transfer learning with data augmentation on IQ-OTH/NCCD lung cancer dataset, retrained on CT-scan images	97.4%
[95] AlexNet	Transfer learning on IQ-OTH/NCCD lung cancer dataset, reclassified into malignant and non-malignant classes	93.54%
[96] GoogleNet	Transfer learning on dataset with malignant and non-malignant classes	94.38%
[97] AlexNet	Feature extraction with hyperparameter tuning, followed by SVM classification for lung cancer classification	98.58%
[97] ResNet50	Feature extraction with hyperparameter tuning, followed by SVM classification for lung cancer classification	96.36%
[98] DECNN	Morphological segmentation and blob feature extractor for lung and liver cancer classification and segmentation, followed by deep ensemble convolutional neural network (DECNN)	75% (dynamic data), 99.8% (static data)
Ours	Without Image Pre-Processing	99.10%
	With Image Pre-Processing	99.84%

Table 4 Comparison of our model with other models in this field

The above table analyses different image classification models that have been performed on the lung tumor detection. As seen from the table that our model has attained highest accuracy compared to others. Without Image pre-processing techniques we attained accuracy of 99.10% and with image pre-processing techniques we attained an accuracy of 99.84%.

This clearly depicts that the model performs well if it gets input data pre-processed like enhancing and de-noising. The final output that has been generated using our model is shown below in figure 27.

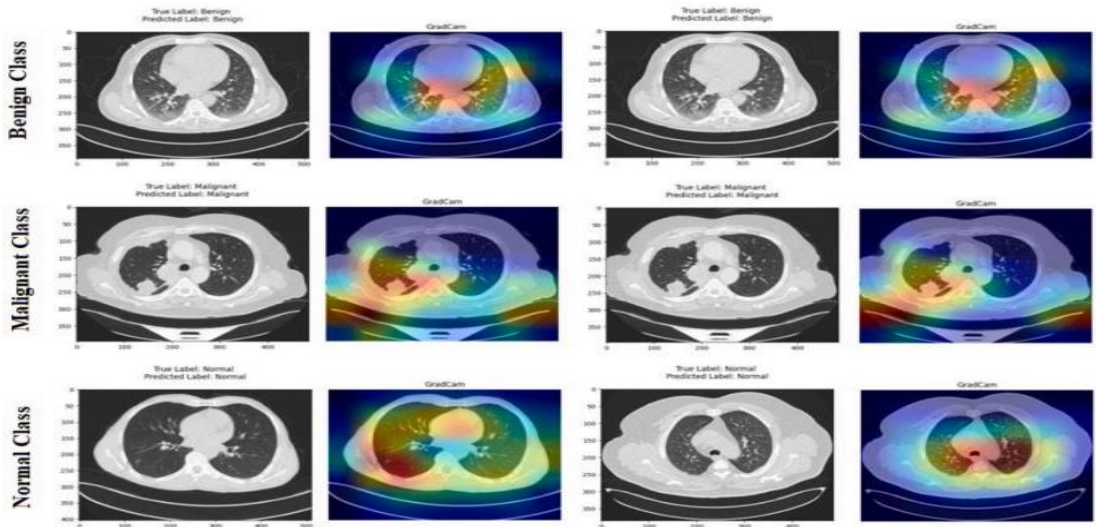


Figure 27 Visualization of our model with class-wise results.

Now, generally classifying and analysing the above output is hard for a normal human being and would seek for doctors' appointment which can a time-consuming process. Keeping this in mind our model did achieve high accuracy, but the main motto of our project is to make it easier for people to understand which saves a lot of cost and time for an individual.

The model that has been discussed is added to our framework and is processed according to the user request and generate possible report form the image. The working is shown in the figure 28 below.

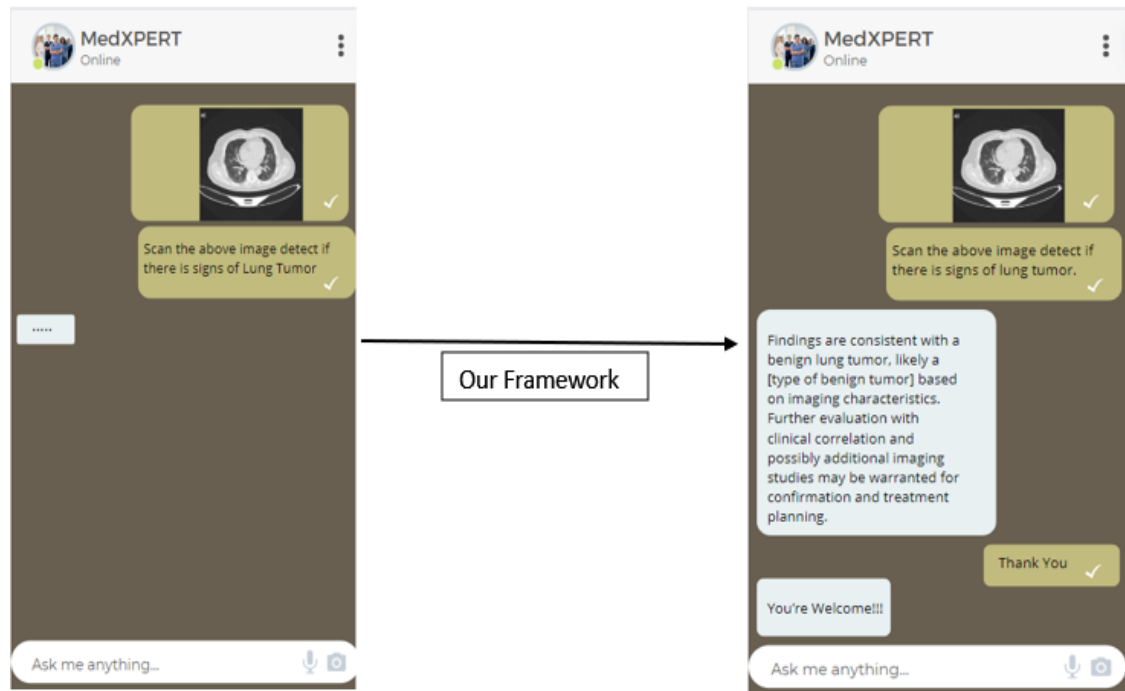


Figure 28 Depicting the proposed interface while working on Lung Tumor detection and generating result

As seen from the above image our framework is a chat interface taking input from user and analysing using the framework and responding accordingly.

7.4 Discussion on Covid-19 Detection

The method that we have proposed by classifying it as two stages which can be clearly understood with the framework built and displayed in the methodology section. The model we proposed is then compared with the existing models that have been performed in the Lung Tumor detection field as shown in table 5.

Our model utilizes foundational models including VGG16, VGG19, Densenet121, InceptionV3, Xception, and Resnet50. Performance evaluation involves using confusion matrices and various metrics such as accuracy, recall, precision, F1-score, loss, and ROC. Among these models, VGG16 demonstrates notably superior performance with a 98.00% accuracy rate. The positive

outcomes of the experiments underscore the effectiveness of the proposed model in detecting and monitoring COVID-19 patients.

Model	Our Models Accuracy		State-of-the-Art Models Accuracy
	Without Image Pre-Processing	With Image Pre-Processing	
Densenet121	97.38%	98.48%	95.77% [99]
Inception V3	93.96%	95.74%	53.4% [100]
Resnet 50	81.29%	88.96%	80% [101]
Xception	94.16%	97.59%	91.95% [99]
VGG16	98%	99.39%	96.18% (Highest) [102]
VGG16	-		76% (Lowest) [101]
VGG19	96.38%	98.64%	96.37% [103]

Table 5 Showing the accuracy comparison of our model with other models

The above table displays all the available models which are used for detecting covid-19 and when they are tested with our methodology different models have achieved different accuracies VGG-16 acquiring the highest and Resnet 50 comparatively lowest and this is for without image pre-processing. And with image pre-processing the accuracies got a little bit higher and making it reliable to refer it.

The models have been trained and tested under many cases and our model which includes the image pre-processing have generated some astonishing results, which is depicted below in figure 29.

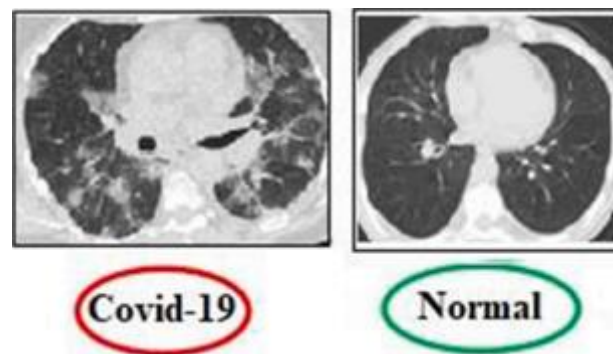


Figure 29 Images generated from our model one identifying Covid-19 and other is completely normal

Our model has generated output as seen above. Now, this can only be accessible to the public if they download and run the code in their devices with their dataset which makes it a hectic and very hard for people who do not know any coding language. This has given our team a major motivation of attaining accessibility of these models to the user. So, we came up with our model MedXPERT which has the framework of the above-mentioned models and

generating a more accurate and reliable information from the data provided by the user.

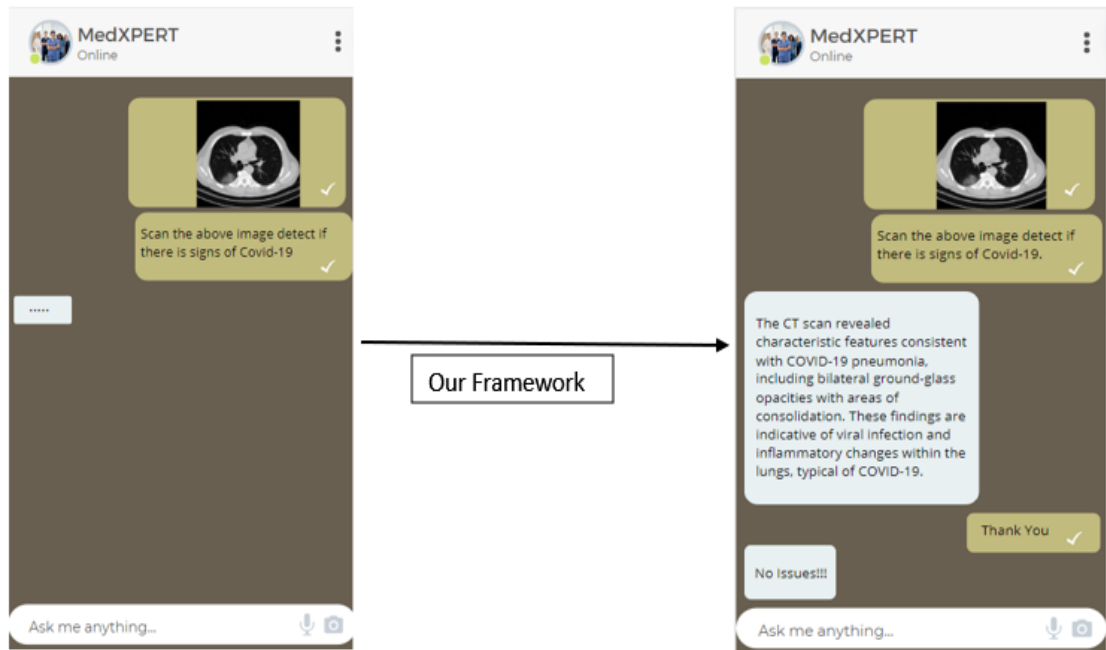


Figure 30 Depicting the proposed Interface while working on Covid-19 detection and generating result

The idea of making it accessible is by ensuring that every human being on earth should access it at their finger-tips. Our main model is a chat interface built using web-technologies. If user has any queries relating to medical, they can come to the open platform and use it. For instance, a user asked the model to check whether there is any sign of covid-19 from a CT image, the image and query are sent to the framework where it gets analysed and generates the possible output to the user in the form of a reply as shown in above figure 30.

The discussions on our proposed models have been done and from all that data. Me and my team can undoubtedly say that our model out-performs all the existing state-of-the art models out in the industries by attaining higher accuracies which gives a sense of belief in people to rely on it, and making it easily accessible for public whenever they are in an emergency.

8. Discussion on Gen-AI in Medical Field

The Outputs are generated in a user-interface which can be seen in this section further down below, this is achieved using OpenAI API KEY and including medical related prompt for more accurate results that the existing GPT-3 models like text-ada-001, text-babbage-001, text-curie-001, and text-davinci-003. Whereas we used GPT-4.

8.1 Gen-AI on Drug Discovery and Development

- Integration of computational methods like molecular docking and dynamics to predict how a drug candidate will interact with its target.
- Increasing use of artificial intelligence (AI) to analyze biological data, identify potential drug candidates, and optimize lead compounds.
- Integration of computational methods like molecular docking and dynamics to predict how a drug candidate will interact with its target.
- Increasing use of artificial intelligence (AI) to analyze biological data, identify potential drug candidates, and optimize lead compounds.

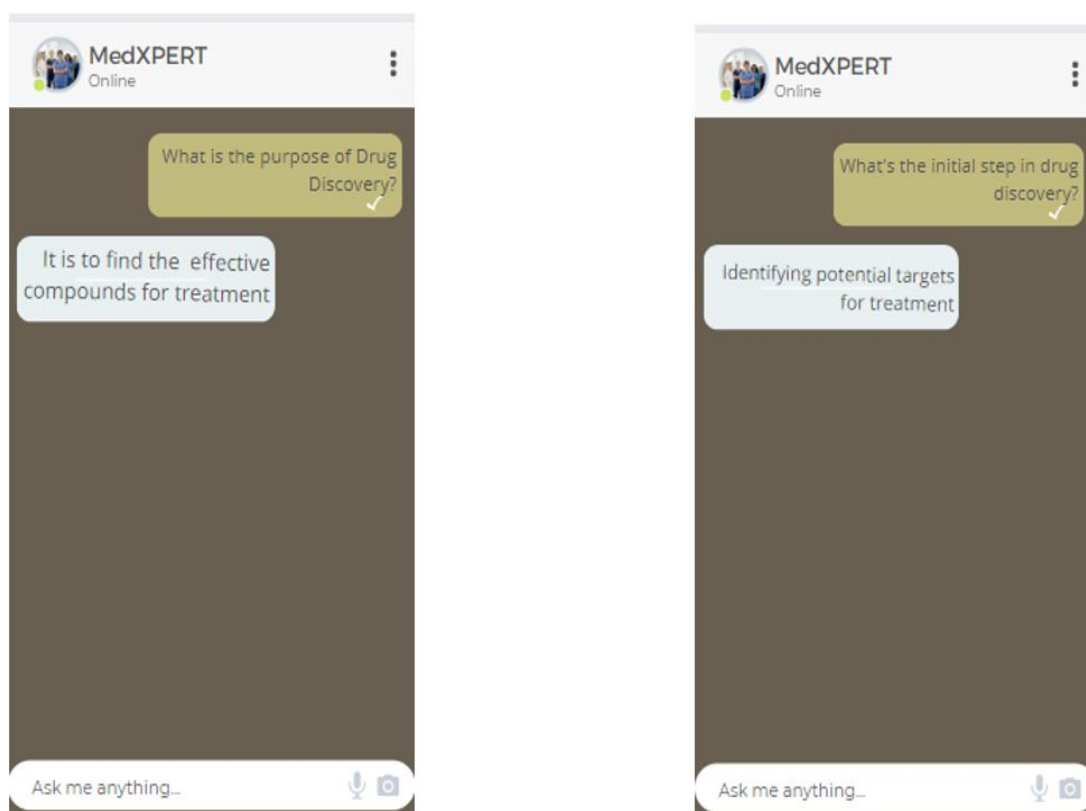


Figure 31 MedXPert Interface

8.2 Gen-AI on Diagnosis

- Integration of AI and machine learning for image analysis and pattern recognition in medical imaging.
- Advancements in personalized medicine based on genetic and molecular profiling.
- Ethical considerations in the use of genetic information for diagnosis and treatment.
- Integration of AI and machine learning for image analysis and pattern recognition in medical imaging.
- Advancements in personalized medicine based on genetic and molecular profiling.
- Ethical considerations in the use of genetic information for diagnosis and treatment.

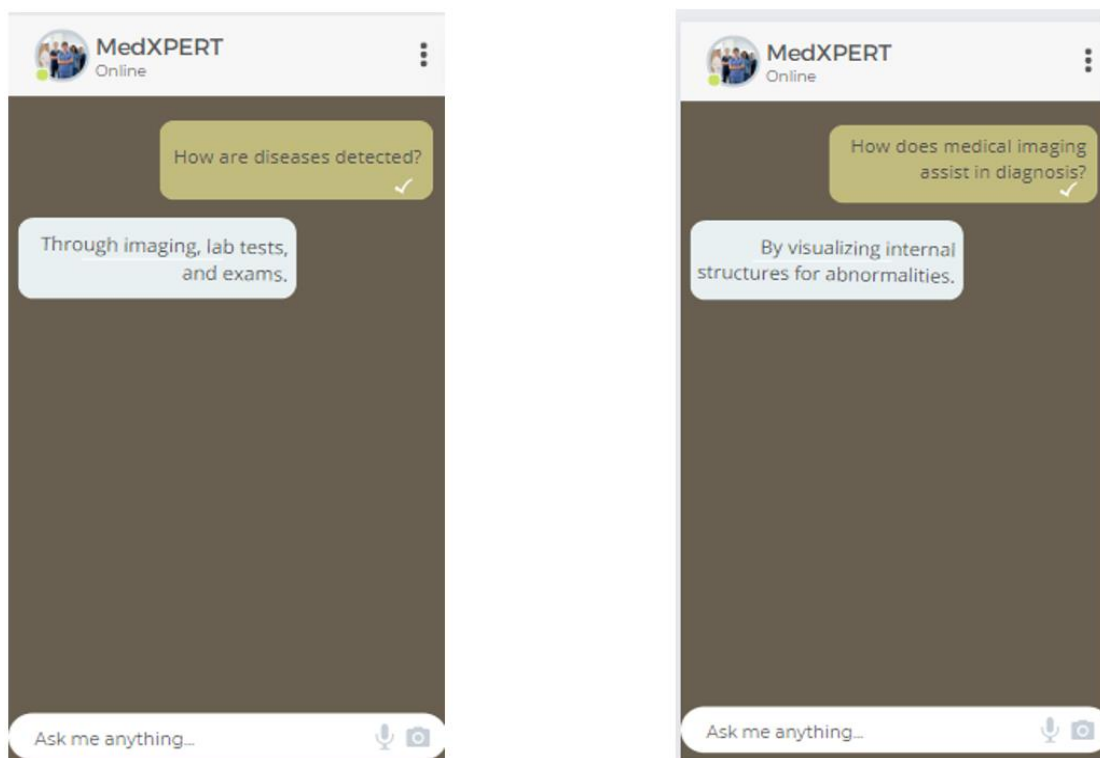


Figure 32 MedXPert Interface

8.3 Gen-AI on Remedies

- Balancing the benefits and risks of different treatment options.
- The rise of precision medicine, tailoring treatments based on individual characteristics.
- Integrative approaches combining conventional and alternative therapies.
- Balancing the benefits and risks of different treatment options.
- The rise of precision medicine, tailoring treatments based on individual characteristics.
- Integrative approaches combining conventional and alternative therapies.

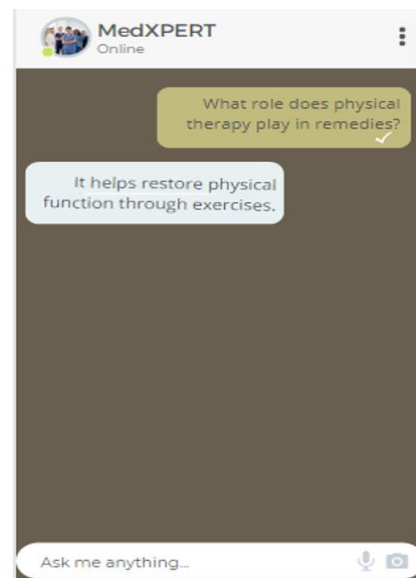
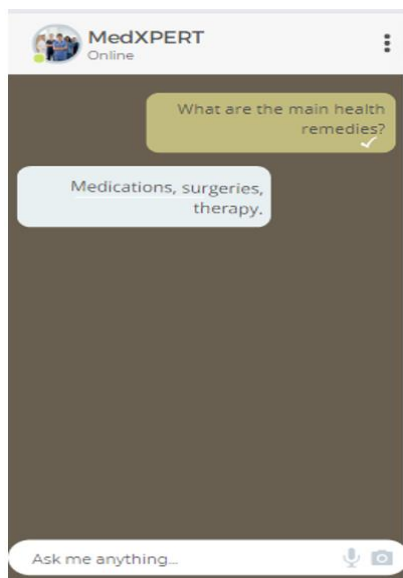


Figure 33 MedXPert Interface

8.4 Gen-AI on Medication (Drug Prediction)

- Challenges in predicting real-world outcomes from in silico models.
- The role of big data and bioinformatics in refining drug prediction models.
- Ethical concerns in using AI for drug discovery.
- The role of big data and bioinformatics in refining drug prediction models.
- Ethical concerns in using AI for drug discovery.

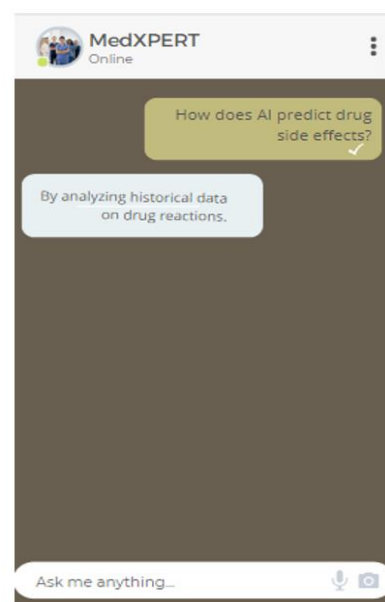


Figure 34 MedXPert Interface

8.5 Virtual Health Care Assistants (Chatbots):

- Privacy and security concerns in handling sensitive health information.
- Improving the emotional intelligence of chatbots for better patient interaction.
- Ensuring proper training and validation to avoid misinformation.
- Privacy and security concerns in handling sensitive health information.
- Improving the emotional intelligence of chatbots for better patient interaction.
- Ensuring proper training and validation to avoid misinformation.

9. Conclusion

The model that we have discussed and proposed in this paper can and will be a game-changer in medical field. Many advancements happened in the medical field such as disease classification, report generation, health evaluations using wearables and many more. These all models have a backbone from which they have derived from and they are the methods like Artificial Intelligence, Machine learning, Deep Learning, Image Pre-Processing techniques and many more.

Those methods are the backbone of my model as well. Apart from those models which are already existing and been worked on, it was time for our group to analyse the most trending methodology of the present decade and we came up with Gen-AI. Now, Gen-AI have made significant role in various fields as discussed before. But now it is time to combine Gen-AI with Medical fields and that was laying foundation of our model MedXPert.

As discussed, there were many models and methodologies that have been worked in medical field for past two decades. So, everything that we had to do was to do unique or go on developing the existing model. So, we incorporated image pre-processing techniques like Image Enhancement, Image De-Noising, Image De-Hazing, Histogram Equalisation and much more.

Many models were referred during the process of generating this model and report. Our proposed methodology for every section that has been discussed was able to beat all the existing models when they are evaluated in terms of accuracy.

The summary of the accuracies that our model was able to generate, **Brain Tumor Detection** - CNN model classification (without image pre-processing): Classification -1: 99.33%, Classification -2: 92.66%, Classification -3: 98.14%. CNN model classification (with image pre-processing): Classification -1: 99.70%, Classification -2: 95.43%, Classification -3: 99.25%. **Alzheimer Detection** - Without Image Pre-Processing: 98.63%, With Image Pre-Processing: 99.42%. **Lung Tumor Detection** - Without Image Pre-Processing: 99.10%, With Image Pre-Processing: 99.84%. **Covid-19 Detection** - Without Image Pre-Processing: Densenet121 - 97.38%, Inception - V3 93.96%, Resnet 50 - 81.29%, Xception - 94.16%, VGG16 - 98%. With Image Pre-Processing: Densenet121 - 98.48%, Inception - V3 - 95.74%, Resnet 50 - 88.96%, Xception - 97.59%, VGG16 - 98.64%.

From the above summary we can really say infusion of Image Pre-processing with the main model framework, really increased the accuracy of their performance. And now, when it comes to Gen-AI there are many Large Language Models present in the industry today and they are also constantly evolving. The Large Language Model that we have used in this model is GPT-4 which is the most efficient Large Language Model present in the market.

Infusing all the models and Gen-AI models have really given birth to our model i.e., MedXPert which as described is a combination of Gen-AI and Medical Field with combining models and making an own personalised medical healthcare chatbot in the finger-tips of human beings.

10. Future Work

The Proposed model MedXPERT, is just the beginning of a new advent in the medical there are lot of new advancements should happen to it, and it should also update accordingly to what is new in the market? And, in the medical field.

Now, the proposed model can only evaluate Brain Tumor, Alzheimer, Lung Tumor, and Covid-19. But the medical field is not only just limited by that they are so many diseases out there for which there is no cure yet, there is no classification technique yet, public who is not aware of it. These all challenges can be resolved by continuous updating of the proposed model soon in the future.

The best way to start updating this model is to incorporate many more features of medical diagnosis techniques like, hard-copy report evaluation, pancreatic cancer, diabetes prediction using retinal images, x-ray classifier and many more. Incorporating all the existing medical diagnostic model to the proposed framework will really improve the performance and many people can rely on it.

The next best way to further upgrade this model is to keep it in a continuous self-updating loop which incorporates the best technologies and methodologies that have been introduced in the market making more efficient and reliable. For example, at present we are using Large Language Model: GPT-4, which is more accurate for now, after 5 or so years another large language model comes with better efficiency. Then this model should analyse the model compare itself with it and then inculcate the efficient model into its framework which can really enhance the ethics of the existing models being truthful to the users.

Now, when it comes to the proposed model itself the model is first planned to develop in a web browser using web technologies, the future work of this model is to create mobile for users to access it anywhere in the globe and try to inculcate it with the health monitoring wearables.

These are most possible future enhancements that can be to the proposed model without losing its reliability, accuracy, robust and user-friendly nature.

REFERENCES

1. Bandi, A.; Adapa, P.V.S.R.; Kuchi, Y.E.V.P.K. The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges. *Future Internet* 2023, 15, 260. <https://doi.org/10.3390/fi15080260>
2. Koohi-Moghadam, M., Bae, K.T. Generative AI in Medical Imaging: Applications, Challenges, and Ethics. *J Med Syst* 47, 94 (2023). <https://doi.org/10.1007/s10916-023-01987-4>
3. Wang, H., Fu, T., Du, Y. et al. Scientific discovery in the age of artificial intelligence. *Nature* 620, 47–60 (2023). <https://doi.org/10.1038/s41586-023-06221-2>
4. Al-Antari, M.A. Artificial Intelligence for Medical Diagnostics – Existing and Future AI Technology! *Diagnostics* 2023, 13, 688. <https://doi.org/10.3390/diagnostics13040688>
5. M. I. Jordan, T. M. Mitchell Machine learning: Trends, perspectives, and prospects. *Science* 349, 255–260 (2015). DOI:10.1126/science.aaa8415
6. Igor Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective, *Artificial Intelligence in Medicine*, Volume 23, Issue 1, 2001, Pages 89–109, ISSN 0933-3657, [https://doi.org/10.1016/S09333657\(01\)00077-X](https://doi.org/10.1016/S09333657(01)00077-X).
7. Qi, Y., Yang, Z., Sun, W. et al. A Comprehensive Overview of Image Enhancement Techniques. *Arch Computat Methods Eng* 29, 583–607 (2022). <https://doi.org/10.1007/s11831-021-09587-6>
8. Leonardo Rundo, Andrea Tangherloni, Marco S. Nobile, Carmelo Militello, Daniela Besozzi, Giancarlo Mauri, Paolo Cazzaniga, MedGA: A novel evolutionary method for image enhancement in medical imaging systems, *Expert Systems with Applications*, Volume 119, 2019, Pages 387–399, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2018.11.013>.
9. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* 521, 436–444 (2015). <https://doi.org/10.1038/nature14539>
10. Bakator, M.; Radosav, D. Deep Learning and Medical Diagnosis: A Review of Literature. *Multimodal Technol. Interact.* 2018, 2, 47. <https://doi.org/10.3390/mti2030047>
11. B. Bayar and M. C. Stamm, "Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection," in *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2691–2706, Nov. 2018, doi: 10.1109/TIFS.2018.2825953.

12. Yadav, S.S., Jadhav, S.M. Deep convolutional neural network based medical image classification for disease diagnosis. *J Big Data* 6, 113 (2019).
<https://doi.org/10.1186/s40537-019-0276-2>
13. Khurana, D., Koli, A., Khatter, K. et al. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl* 82, 3713–3744 (2023).
<https://doi.org/10.1007/s11042-022-13428-4>
14. Stéphane Meystre, Peter J. Haug, Natural language processing to extract medical problems from electronic clinical documents: Performance evaluation, *Journal of Biomedical Informatics*, Volume 39, Issue 6, 2006, Pages 589-599, ISSN 1532-0464,
<https://doi.org/10.1016/j.jbi.2005.11.004>.
15. Xu, Y., Wei, H., Lin, M. et al. Transformers in computational visual media: A survey. *Comp. Visual Media* 8, 33–62 (2022). <https://doi.org/10.1007/s41095-021-0247-3>
16. Kelei He, Chen Gan, Zhuoyuan Li, Islem Rekik, Zihao Yin, Wen Ji, Yang Gao, Qian Wang, Junfeng Zhang, Dinggang Shen, Transformers in medical image analysis, *Intelligent Medicine*, Volume 3, Issue 1, 2023, Pages 59-78, ISSN 2667-1026,
<https://doi.org/10.1016/j.imed.2022.07.002>.
17. A. Dash, J. Ye and G. Wang, "A Review of Generative Adversarial Networks (GANs) and Its Applications in a Wide Variety of Disciplines: From Medical to Remote Sensing," in *IEEE Access*, vol. 12, pp. 18330-18357, 2024, doi: 10.1109/ACCESS.2023.3346273.
18. Paladugu, P.S., Ong, J., Nelson, N. et al. Generative Adversarial Networks in Medicine: Important Considerations for this Emerging Innovation in Artificial Intelligence. *Ann Biomed Eng* 51, 2130–2142 (2023). <https://doi.org/10.1007/s10439-023-03304-z>
19. Alberts, I., Mercolli, L., Pyka, T. et al. Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be?. *Eur J Nucl Med Mol Imaging* 50, 1549–1552 (2023). <https://doi.org/10.1007/s00259-023-06172-w>
20. Jiancheng Yang, Hongwei Bran Li, Donglai Wei, The impact of ChatGPT and LLMs on medical imaging stakeholders: Perspectives and use cases, *Meta-Radiology*, Volume 1, Issue 1, 2023, 100007, ISSN 2950-1628,
<https://doi.org/10.1016/j.metrad.2023.100007>.
21. Claudia Vitolo, Yehia Elkhatib, Dominik Reusser, Christopher J.A. Macleod, Wouter Buytaert, Web technologies for environmental Big Data, *Environmental Modelling & Software*, Volume 63, 2015, Pages 185-198, ISSN 1364-8152,
<https://doi.org/10.1016/j.envsoft.2014.10.007>.

22. Chirumamilla S, Gulati M. Patient Education and Engagement through Social Media. *Curr Cardiol Rev.* 2021;17(2):137-143. doi: 10.2174/1573403X15666191120115107. PMID: 31752656; PMCID: PMC8226210.
23. Aguirre RR, Suarez O, Fuentes M, Sanchez-Gonzalez MA. Electronic Health Record Implementation: A Review of Resources and Tools. *Cureus.* 2019 Sep 13;11(9):e5649. doi: 10.7759/cureus.5649. PMID: 31700751; PMCID: PMC6822893.
24. Lardo AC. Real-time magnetic resonance imaging: diagnostic and interventional applications. *Pediatr Cardiol.* 2000 Jan-Feb;21(1):80-98. doi: 10.1007/s002469910010. PMID: 10672617
25. Bodduluri S, Reinhardt JM, Hoffman EA, Newell JD Jr, Bhatt SP. Recent Advances in Computed Tomography Imaging in Chronic Obstructive Pulmonary Disease. *Ann Am Thorac Soc.* 2018 Mar;15(3):281-289. doi: 10.1513/AnnalsATS.201705-377FR. PMID: 28812906; PMCID: PMC5880521.
26. Cross AJ, Elliott RA, Petrie K, Kuruvilla L, George J. Interventions for improving medication-taking ability and adherence in older adults prescribed multiple medications. *Cochrane Database Syst Rev.* 2020 May 8;5(5):CD012419. doi: 10.1002/14651858.CD012419.pub2. PMID: 32383493; PMCID: PMC7207012.
27. Bernardo F, Zbyszyński M, Grierson M, Fiebrink R. Designing and Evaluating the Usability of a Machine Learning API for Rapid Prototyping Music Technology. *Front Artif Intell.* 2020 Apr 3;3:13. doi: 10.3389/frai.2020.00013. PMID: 33733132; PMCID: PMC7861239.
28. Amin, J., Sharif, M., Haldorai, A. et al. Brain tumor detection and classification using machine learning: a comprehensive survey. *Complex Intell. Syst.* 8, 3161–3183 (2022). <https://doi.org/10.1007/s40747-021-00563-y>
29. Badža MM, Barjaktarović MC (2020) Classification of brain tumors from MRI images using a convolutional neural network. *Appl Sci* 10(6):1–13. <https://doi.org/10.3390/app10061999>
30. Saeedi, S., Rezayi, S., Keshavarz, H. et al. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Med Inform Decis Mak* 23, 16 (2023). <https://doi.org/10.1186/s12911-023-02114-6>
31. Diyuan Lu, Nenad Polomac, Iskra Gacheva, Elke Hattingen, Jochen Triesch, Human-Expert-Level Brain Tumor Detection Using Deep Learning with Data Distillation and Augmentation, arXiv:2006.12285, <https://doi.org/10.48550/arXiv.2006.12285>

32. Zachary Schwehr and Sriman Achanta, Brain Tumor Segmentation Based on Deep Learning, Attention Mechanisms, and Energy-Based Uncertainty Prediction (2023), Image and Video Processing, arXiv:2401.00587, <https://doi.org/10.48550/arXiv.2401.00587>
33. Ranjbarzadeh, R., Bagherian Kargari, A., Jafarzadeh Ghouschi, S. et al. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. Sci Rep 11, 10930 (2021). <https://doi.org/10.1038/s41598-021-90428-8>
34. Tran, K.A., Kondrashova, O., Bradley, A. et al. Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med 13, 152 (2021). <https://doi.org/10.1186/s13073-021-00968-x>
35. N. Amoroso, D. Diacono, A. Fanizzi, M. La Rocca, A. Monaco, A. Lombardi, C. Guaragnella, R. Bellotti, and S. Tangaro, "Deep learning reveals Alzheimer's disease onset in MCI subjects: Results from an international challenge," J. Neurosci. Methods, vol. 302, pp. 3-9, May 2018.
36. Qiu, S., Miller, M.I., Joshi, P.S. et al. Multimodal deep learning for Alzheimer's disease dementia assessment. Nat Commun 13, 3404 (2022). <https://doi.org/10.1038/s41467-022-31037-5>
37. Diogo, V.S., Ferreira, H.A., Prata, D. et al. Early diagnosis of Alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach. Alz Res Therapy 14, 107 (2022). <https://doi.org/10.1186/s13195-022-01047-y>
38. Behrendt, F., Bengs, M., Bhattacharya, D. et al. A systematic approach to deep learning-based nodule detection in chest radiographs. Sci Rep 13, 10120 (2023). <https://doi.org/10.1038/s41598-023-37270-2>
39. A. . Gudur, H. . Sivaraman, and V. . Vimal, "Deep Learning-Based Detection of Lung Nodules in CT Scans for Cancer Screening", Int J Intell Syst Appl Eng, vol. 11, no. 7s, pp. 20-28, Jul. 2023.
40. Detecting COVID-19 and Community Acquired Pneumonia using Chest CT scan images with Deep Learning, Shubham Chaudhary and Sadbhawna and Vinit Jakheti and Badri N Subudhi and Ujjwal Baid and Sharath Chandra Guntuku, 2021, arXiv:2104.05121, <https://doi.org/10.48550/arXiv.2104.05121>
41. E. Dong, H. Du, and L. Gardner, "An interactive webbased dashboard to track covid-19 in real time," The Lancet infectious diseases, vol. 20, no. 5, pp. 533-534, 2020.

42. F. Grillet, J. Behr, P. Calame, S. Aubry, and journal=Radiology volume=296 number=3 pages=E186–E188 year=2020 publisher=Radiological Society of North America Delabrousse, E., “Acute pulmonary embolism associated with covid-19 pneumonia detected with pulmonary ct angiography,” .
43. C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and X. Wang, “Deep learning-based detection for covid-19 from chest ct using weak label,” MedRxiv, 2020.
44. P. Afshar, S. Heidarian, N. Enshaei, F. Naderkhani, R. Moezedin J., A. Oikonomou, F. Fard, K. Samimi, K. Plataniotis, and A. Mohammadi, “Covid-ct-md: Covid-19 computed tomography (ct) scan dataset applicable in machine learning and deep learning,” arXiv, 09 2020.
45. S. A. Harmon, T. H. Sanford, S. Xu, E. B. Turkbey, H. Roth, ..., M. Cariati, G. Carrafiello, P. An, B. J. Wood, and B. Turkbey, “Artificial intelligence for the detection of covid-19 pneumonia on chest ct using multinational datasets,” Nature Communications, vol. 11, pp. 2041–1723, 2020.
46. Zhao, W., Jiang, W. & Qiu, X. Deep learning for COVID-19 detection based on CT images. Sci Rep 11, 14353 (2021). <https://doi.org/10.1038/s41598-021-93832-2>
47. Chen, Z., Pawar, K., Ekanayake, M. et al. Deep Learning for Image Enhancement and Correction in Magnetic Resonance Imaging – State-of-the-Art and Challenges. J Digit Imaging 36, 204–230 (2023). <https://doi.org/10.1007/s10278-022-00721-9>
48. Rasheed, Z.; Ma, Y.-K.; Ullah, I.; Ghadi, Y.Y.; Khan, M.Z.; Khan, M.A.; Abdusalomov, A.; Alqahtani, F.; Shehata, A.M. Brain Tumor Classification from MRI Using Image Enhancement and Convolutional Neural Network Techniques. Brain Sci. 2023, 13, 1320. <https://doi.org/10.3390/brainsci13091320>
49. Xiao Z, Ding Y, Lan T et al (2017) Brain MR image classification for alzheimer’s disease diagnosis based on multifeature fusion. Comput Math Methods Med 2017:1952373. <https://doi.org/10.1155/2017/1952373>
50. Shakeel, P.M., Burhanuddin, M.A. & Desa, M.I. Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Comput & Applic 34, 9579–9592 (2022). <https://doi.org/10.1007/s00521-020-04842-6>
51. Goyal, S., Singh, R. Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques. J Ambient Intell Human Comput 14, 3239–3259 (2023). <https://doi.org/10.1007/s12652-021-03464-7>
52. Conde, Marcos & Geigle, Gregor & Timofte, Radu. (2024). High-Quality Image Restoration Following Human Instructions. 10.48550/arXiv.2401.16468.

53. PromptMRG: Diagnosis-Driven Prompts for Medical Report Generation, Haibo Jin and Haoxuan Che and Yi Lin and Hao Chen, 2024, arXiv:2308.12604, <https://doi.org/10.48550/arXiv.2308.12604>
54. Multi-modal Pre-training for Medical Vision-language Understanding and Generation: An Empirical Study with A New Benchmark, Li Xu and Bo Liu and Ameer Hamza Khan and Lu Fan and Xiao-Ming Wu, 2023, arXiv:2306.06494, <https://doi.org/10.48550/arXiv.2306.06494>
55. Ya Y, Ji L, Jia Y, Zou N, Jiang Z, Yin H, Mao C, Luo W, Wang E and Fan G (2022) Machine Learning Models for Diagnosis of Parkinson's Disease Using Multiple Structural Magnetic Resonance Imaging Features. *Front. Aging Neurosci.* 14:808520. doi: 10.3389/fnagi.2022.808520
56. Dritsas, E.; Trigka, M. Stroke Risk Prediction with Machine Learning Techniques. *Sensors* 2022, 22, 4670. <https://doi.org/10.3390/s22134670>
57. Artificial Intelligence in Drug Discovery: Applications and Techniques Jianyuan Deng, Zhibo Yang, Iwao Ojima, Dimitris Samaras, and Fusheng Wang.
58. Generative chemistry: drug discovery with deep learning generative models Yuemin Bian^{1,2} and Xiang-Qun Xie^{1,2,3,4}.
59. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials Maria Bordukova, Nikita Makarov, Raul Rodriguez-Esteban, Fabian Schmich & Michael P. Menden.
60. Generative AI in Medicine and Healthcare: Promises, Opportunities and Challenges Peng Zhang ¹ and Maged N. Kamel Boulos.
61. Will Artificial Intelligence for Drug Discovery Impact Clinical Pharmacology? Alex Zhavoronkov, Quentin Vanhaelen and Tudor I. Oprea.
62. Deep generative molecular design reshapes drug discovery, Xiangxiang Zeng,¹ Fei Wang,² Yuan Luo,³ Seung-gu Kang,⁴ Jian Tang,⁵ Felice C. Lightstone,⁶ Evandro F. Fang,^{7,8} Wendy Cornell,⁴ Ruth Nussinov,^{9,10} and Feixiong Cheng).
63. Generative artificial intelligence in drug discovery: basic framework, recent advances, challenges, and opportunities Amit Gangwal, Azim Ansari, Iqrar Ahmad, Abul Kalam Azad, Vinoth Kumarasamy, Vetriselvan Subramaniyan, and Ling Shing Wong.
64. Artificial Intelligence for Drug Discovery: Are We There Yet? Catrin Hasselgren¹ and Tudor I. Oprea.

65. Generative Artificial Intelligence in (laboratory) medicine: friend or foe? Davide Negrini, Giuseppe Lippi.
66. ChatGPT: The transformative influence of generative AI on science and healthcare Julian Varghese, Julius Chapiro.
67. Johnson DR, Guerin JB, Giannini C, Morris JM, Eckel LJ, Kaufmann TJ (2017) 2016 updates to the WHO brain tumor classification system: what the radiologist needs to know. *Radiographics* 37:2164–2180
68. Koon-Pong Wong, Dagan Feng, S. R. Meikle and M. J. Fulham, "Segmentation of dynamic PET images using cluster analysis," in *IEEE Transactions on Nuclear Science*, vol. 49, no. 1, pp. 200-207, Feb. 2002, doi: 10.1109/TNS.2002.998752.
69. MRI can assess structures that may be difficult to discern in a CT scan, offering superior contrast for soft tissues and enhancing the clarity of anatomical structures. Hess CP, Purcell D (2012) Exploring the brain: Is CT or MRI better for brain imaging. *UCSF Dep Radiol Biomed Imaging* 11:1–11
70. Saad NM, Bakar SARSA, Muda AS, Mokji MM (2015) Review of brain lesion detection and classification using neuroimaging analysis techniques. *J Teknol* 74:1–13
71. F. M. J. M. Shamrat et al., "AlzheimerNet: An Effective Deep Learning Based Proposition for Alzheimer's Disease Stages Classification From Functional Brain Changes in Magnetic Resonance Images," in *IEEE Access*, vol. 11, pp. 16376-16395, 2023, doi:10.1109/ACCESS.2023.3244952.
72. Raza, S., Schwartz, B. Constructing a disease database and using natural language processing to capture and standardize free text clinical information. *Sci Rep* 13, 8591 (2023). <https://doi.org/10.1038/s41598-023-35482-0>
73. Pawar K, Chen Z, Shah NJ, Egan GF. Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation. *NMR Biomed.* 2019; 35(4):e4225. <https://doi.org/10.1002/nbm.4225>
74. Hongjun An, Hyeong-Geol Shin, Sooyeon Ji, Woojin Jung, Sehong Oh, Dongmyung Shin, Juhyung Park, Jongho Lee, DeepResp: Deep learning solution for respiration-induced B0 fluctuation artifacts in multi-slice GRE, *NeuroImage*, Volume 224, 2021, 117432, ISSN 1053-8119, <https://doi.org/10.1016/j.neuroimage.2020.117432>.
75. Vishal Venkatesh, Neeraj Sharma, Munendra Singh, Intensity inhomogeneity correction of MRI images using InhomoNet, *Computerized Medical Imaging and Graphics*, Volume 84, 2020, 101748, ISSN 0895-6111, <https://doi.org/10.1016/j.compmedimag.2020.101748>.

76. Blake E. Dewey, Can Zhao, Jacob C. Reinhold, Aaron Carass, Kathryn C. Fitzgerald, Elias S. Sotirchos, Shiv Saidha, Jiwon Oh, Dzung L. Pham, Peter A. Calabresi, Peter C.M. van Zijl, Jerry L. Prince, DeepHarmony: A deep learning approach to contrast harmonization across scanner changes, *Magnetic Resonance Imaging*, Volume 64, 2019, Pages 160-170, ISSN 0730-725X, <https://doi.org/10.1016/j.mri.2019.05.041>.
77. Maosong Ran, Jinrong Hu, Yang Chen, Hu Chen, Huaqiang Sun, Jiliu Zhou, Yi Zhang, Denoising of 3D magnetic resonance images using a residual encoder-decoder Wasserstein generative adversarial network, *Medical Image Analysis*, Volume 55, 2019, Pages 165-180, ISSN 1361-8415, <https://doi.org/10.1016/j.media.2019.05.001>.
78. Yu Qin, Zhiwen Liu, Chenghao Liu, Yuxing Li, Xiangzhu Zeng, Chuyang Ye, Super-Resolved q-Space deep learning with uncertainty quantification, *Medical Image Analysis*, Volume 67, 2021, 101885, ISSN 1361-8415, <https://doi.org/10.1016/j.media.2020.101885>.
79. Mzoughi H, Njeh I, Wali A, Slima M, Ben BenHamida A, Mhiri C, Mahfoudhe K (2020) Deep Multi-Scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification. *J Digit Imag* 33(4):903–915. <https://doi.org/10.1007/s10278-020-00347-9>
80. Ayadi W, Elhamzi W, Charfi I, Atri M (2021) Deep CNN for brain tumor classification. *Neural Process Lett* 53(1):671–700. <https://doi.org/10.1007/s11063-020-10398-2>
81. Pereira S, Meier R, Alves V, Reyes M, Silva CA (2018) Automatic brain tumor grading from MRI data using convolutional neural networks and quality assessment. *Understanding and interpreting machine learning in medical image computing applications*. Springer, Cham, pp 106–114
82. Talo M, Baloglu UB, Yıldırım Ö, Rajendra Acharya U (2019) Application of deep transfer learning for automated brain abnormality classification using MR images. *Cogn Syst Res* 54(12):176–188. <https://doi.org/10.1016/j.cogsys.2018.12.007>
83. Abiwinanda N, Hanif M, Hesaputra ST, Handayani A, Mengko TR. Brain tumor classification using convolutional neural network. In: *World congress on medical physics and biomedical engineering 2018*. Singapore: Springer; 2019. p. 183–9.
84. Abd-Ellah, M., Awad, A., Khalaf, A. et al. Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using convolutional neural networks. *J Image Video Proc*. 2018, 97 (2018). <https://doi.org/10.1186/s13640-018-0332-4>.

85. Rehman, A., Naz, S., Razzak, M.I. et al. A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning. *Circuits Syst Signal Process* 39, 757–775 (2020). <https://doi.org/10.1007/s00034-019-01246-3>
86. A. Khvostikov, K. Aderghal, J. Benois-Pineau, A. Krylov, and G. Catheline, “3D CNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies,” 2018, arXiv:1801.05968.
87. J. Islam and Y. Zhang, “Understanding 3D CNN behavior for Alzheimer’s disease diagnosis from brain PET scan,” 2019, arXiv:1912.04563.
88. M. Amini, M. Pedram, A. Moradi, and M. Ouchani, “Diagnosis of Alzheimer’s disease severity with fMRI images using robust multitask feature extraction method and convolutional neural network (CNN),” *Comput. Math. Methods Med.*, vol. 2021, pp. 1–15, Apr. 2021.
89. A. Ebrahimi and S. Luo, “Convolutional neural networks for Alzheimer’s disease detection on MRI images,” *J. Med. Imag.*, vol. 8, no. 2, 2021, Art. no. 24503.
90. Basaia S, Agosta F, Wagner L et al (2019) Automated classification of Alzheimer’s disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage Clin* 21:101645. <https://doi.org/10.1016/j.nicl.2018.101645>
91. Liu M, Li F, Yan H et al (2020) A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer’s disease. *Neuroimage* 208:116459. <https://doi.org/10.1016/j.neuroimage.2019.116459>
92. Lu X, Wu H, Zeng Y (2019) Classification of Alzheimer’s disease in MobileNet. *J Phys Conf Ser.* <https://doi.org/10.1088/1742-6596/1345/4/042012>
93. Shen T, Jiang J, Lu J et al (2019) Predicting alzheimer disease from mild cognitive impairment with a deep belief network based on 18F-FDG-PET images. *Mol Imag* 18:1–9. <https://doi.org/10.1177/1536012119877285>
94. Humayun, M.; Sujatha, R.; Almuayqil, S.N.; Jhanjhi, N.Z. A Transfer Learning Approach with a Convolutional Neural Network for the Classification of Lung Carcinoma. *Healthcare* 2022, 10, 1058. <https://doi.org/10.3390/healthcare10061058>
95. Hamdalla F. Al-Yasriy et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 928 022035 DOI 10.1088/1757-899X/928/2/022035
96. AL-Huseiny M.S., Sajit A.S., Transfer learning with GoogLeNet for detection of lung cancer *Indonesian J. Electr. Eng. Comput. Sci.*, 22 (2021), pp. 1078-1086 DOI: <https://doi.org/10.11591/ijeecs.v22.i2.pp1078-1086>

97. NARİN, D., & ONUR, T. Ö. (2022). The Effect of Hyper Parameters on the Classification of Lung Cancer Images Using Deep Learning Methods. *Erzincan University Journal of Science and Technology*, 15(1), 258-268.
<https://doi.org/10.18185/erzifbed.1006560>
98. Raza, R.; Zulfiqar, F.; Tariq, S.; Anwar, G.B.; Sargano, A.B.; Habib, Z. Melanoma Classification from Dermoscopy Images Using Ensemble of Convolutional Neural Networks. *Mathematics* 2022, 10, 26. <https://doi.org/10.3390/math10010026>
99. Lahsaini I, Daho MEH, Chikh MA. Deep transfer learning based classification model for covid-19 using chest CT-scans. *Pattern Recognition Lett* 2021;152:122–8.
100. Ebenezer AS, Kanmani SD, Sivakumar M, Priya SJ. Effect of image transformation on EfficientNet model for COVID-19 CT image classification. *Mater Today: Proceedings* 2022;51:2512–9.
101. He X, Yang X, Zhang S, Zhao J, Zhang Y, Xing E, Xie P. Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. *medrxiv* 2020.
102. Shaik NS, Cherukuri TK. Transfer learning based novel ensemble classifier for COVID-19 detection from chest CT-scans. *Comput Biol Med* 2022;141:105127.
103. Jangam E, Barreto AAD, Annavarapu CSR. Automatic detection of COVID-19 from chest CT scan and chest X-Rays images using deep learning, transfer learning and stacking. *Applied Intell* 2022;52(2):2243–59.