

# AI-Assisted Early Diagnosis of Alzheimer's Disease: A Clinical Evaluation of Machine and Deep Learning Models

Safdar Sardar Khan<sup>1\*</sup>, Rijvan Beg<sup>2</sup>, Shredha Parmar<sup>1</sup>,  
Mohammed Irfan Khan<sup>1</sup>, Arpit Deo<sup>1</sup> and Sadaf Akhtar<sup>1</sup>

<sup>1</sup>Department of Computer Science and Engineering, Medicaps University, Indore MP, India.

<sup>2</sup>Department of Computer Science and Engineering, SRM University, Amravati,  
Andhra Pradesh, India.

\*Corresponding Author E-mail: safdar.cse@gmail.com

<https://dx.doi.org/10.13005/bpj/3411>

(Received: 21 May 2025; accepted: 14 October 2025)

This study aims to evaluate the clinical efficacy of machine learning and deep learning models, particularly EfficientNetB7, for the early diagnosis of Alzheimer's disease using neuroimaging and clinical data. Machine learning and deep learning approaches to evaluate complex data, including neuroimaging and clinical information, aiding in the timely detection of Alzheimer's disease and delivering precise diagnoses. Machine learning algorithms, through comprehensive analysis of diverse data sources, have the capacity to discern intricate patterns and features that may elude human observation. Through training on extensive datasets, these models achieve heightened accuracy and improved generalization in diagnosing Alzheimer's disease. Deep learning, a subset of machine learning characterized by neural networks with numerous layers, enhances the capacity to decipher and extract nuanced patterns from extensive and intricate datasets. The development of reliable machine learning and deep learning models holds significant potential for enhancing early detection, a pivotal factor for timely interventions and appropriate treatment plans. Early diagnosis can lead to more effective disease management, maybe improving the standard of life for those suffering from Alzheimer's disease and reducing the strain on the healthcare system and carers. In a performance evaluation, the EfficientNetB7 model demonstrated exceptional accuracy, achieving a score of 99.36%. This underscores its superior performance in comparison to other models. Furthermore, it displayed a validation accuracy of 49.48%, showcasing its robustness and capability. The model maintained a validation loss of 2.9, while the training loss was impressively low at 0.017. When benchmarked against other machine learning models, EfficientNetB7 consistently outperforms, underscoring its potential for elevating diagnostic accuracy in Alzheimer's disease detection.

**Keywords:** Alzheimer's disease detection; Deep learning; EfficientNetB7;  
Health care; Machine learning.

---

Previous studies have demonstrated practical applications of AI in Alzheimer's detection. For instance, Yao et al. <sup>1</sup> applied AI to brain MRI images in a clinical cohort of 150 patients, achieving 85% accuracy in distinguishing early-stage Alzheimer's from healthy controls, highlighting its potential in routine neuroimaging

workflows. Similarly, Hu et al. <sup>2</sup> conducted a meta-analysis of MRI-based ML models across 20 studies involving over 5,000 participants, showing improved diagnostic precision in hospital settings compared to traditional cognitive tests. These cases underscore how AI can integrate with clinical practices to enable earlier interventions, such as

in memory clinics where timely diagnosis reduces caregiver burden.

Memory loss is one of the cognitive skills that is consistently worsened by Alzheimer's disease, a kind of dementia, intellectual capacity, behavioral control, and physical skills. Midway through a patient's 60s is when they often report their first instance of late-onset symptoms, on average. It is quite unusual for Alzheimer's disease to present itself in people under the age of 60 years and 10 months. The leading cause of dementia among seniors is unquestionably Alzheimer's disease. The disease was given the name Alzheimer in recognition of Dr. Alois Alzheimer. Dr. Alzheimer conducted an autopsy on a woman whose death in 1906 was attributed to an unidentified mental condition. He observed anomalies in her brain. She suffered from amnesia, difficulties speaking, and erratic behavior as a result of her condition. Following her passing, he carried out an autopsy on her brain and found numerous amyloid plaques and neurofibrillary (or tau) tangles. Now understood to be disorders linked to Alzheimer's disease are amyloid plaques & neurofibrillary (or tau) tangles.<sup>2</sup> Plaques and tangles are still a primary indicator of Alzheimer's disease, as they are seen throughout the brain in affected individuals. Another feature is the breakdown of synapses, which are connections between neurons. Information is transmitted throughout the brain and out to the muscles and organs of the body by neurons. There is speculation that a number of other complicated disorders in the brain have a role in the progression of Alzheimer's disease. The parts of the brain that are responsible for memory, such as the entorhinal cortex and the hippocampus, are among the first to be damaged. Affected are parts of the cortex that participate in subsequent development, such as those engaged in language, thinking, and social interaction. Damage is also done to a great deal of the remainder of the brain's areas. Memory loss is frequently one of the earliest indications of Alzheimer's disease, despite the fact that the progression of the disease manifests itself differently in each individual. A slowing of reasoning or judgment, issues with eyesight or spatial awareness, and an inability to find the right words could all be early indicators of Alzheimer's disease. Even though moderate cognitive impairment (MCI) is not always a

precursor to Alzheimer's condition, it is probable that some people who have it will eventually develop the disease.<sup>3-5</sup>

Alzheimer's patients commonly struggle with "the little tasks" of daily life, such driving and cooking, and paying their bills. They might ask the same questions over and over again, they might forget where they placed things, and they might struggle to understand the simplest of ideas. As the disease progresses, some people may experience rising levels of anxiety, rage, and even violent behavior. Currently, Alzheimer's disease is the sixth most common cause of mortality in the United States, albeit the interval between diagnosis and passing away varies. If the patient is less than 80 years old at the time of diagnosis, their life expectancy may be as high as ten years or more, but it may be as low as three or four years if they are older than 80 years old. In recent years, significant headway has been made in the investigation and clinical testing of potential Alzheimer's disease therapy; yet, the condition still does not have a cure. The Food and Drug Administration has given its approval for the use of many medications in the treatment of Alzheimer's disease.<sup>6-8</sup>

It is still unclear what triggers Alzheimer's disease. Numerous genetic and environmental variables have been linked to its emergence. The most important genetic component is an APOE allele. Additional risk factors include having a history of head trauma, high blood pressure, and severe depression. The illness process is characterized by the breakdown of neuronal connections and the development of amyloid plaques in the brain. After a diagnosis has been determined based on the patient's medical history and the findings of cognitive testing, medical imaging and blood tests are utilized to rule out any further potential reasons. The initial symptoms are quickly written off as the product of normal brain aging. Brain tissue must be examined postmortem in order to provide a definitive diagnosis. As one becomes older, keeping meaningful social connections, maintaining an active lifestyle, and adhering to a nutritious diet might help reduce the risk of cognitive decline including Alzheimer's disease. At this time, there is no treatment or remedy currently available that can halt or decrease its spread. Individuals affected become overly dependent on their caretakers as they rely on them

for more and more help. Social, psychological, bodily, and economic factors are all potential sources of stress. By engaging in a fitness program, daily chores may provide better results. Despite the fact that there is little benefit and an elevated risk of premature mortality, antipsychotics are frequently given for dementia patients who are demonstrating behavioral problems or psychosis. There are millions of people who are affected with AD worldwide. This degenerative disease's symptoms include memory loss, mental decline, and changed behavior. Research into biomarkers is a powerful weapon in the war against AD. Researchers can now accurately diagnose individuals with Alzheimer's disease and observe their disease evolve through several phases using this cutting-edge technology. Other neurological impairments that may be linked to the disease can be identified utilizing biomarkers. Identifying biomarkers for Alzheimer's disease is an important step toward understanding the disease and creating improved treatments for it. This article will discuss the significance of biomarker-based research in the fight against Alzheimer's disease. Various biomarkers can be used to diagnose various neurological disorders. Abnormal behavior in Alzheimer's patients has been linked to a disease in the brain caused by these mutations. Biological reasoning may be used to the little protein known as the Amyloid Precursor Protein (APP), which is present in several brain regions, to eliminate potential sites. Amyloid precursor protein (APP) aids in the regeneration, degradation, and recycling of damaged neurons over the course of certain diseases. Deficiencies in APP can increase synaptic proteases, which can produce neuro region deficits and disruptions in the brain. Your brain's synapses are what carry signals between neurons. These neurons are shielded from environmental influences by a membrane made of APP. Therefore, groups of proteases found outside the blood arteries can cause inflammation. Inflammation is associated with numerous symptoms and progressions of Alzheimer's disease (AD).<sup>9</sup> Clinical biomarkers, genetic biomarkers, positron emission tomography (PET), biospecimen biomarkers, structural magnetic resonance imaging (sMRI), cerebrospinal fluid (CSF), positron emission tomography (PET), fluid attenuated inversion recovery (FLAIR) imaging, diffusion tensor imaging (DTI), and

electroencephalography (EEG) can all provide insight into the efficiency levels of specific brain regions. Medical professionals and academics rely on these methods for diagnosing Alzheimer's disease and its many variants. These biomarkers can help doctors detect conditions affecting the brain and nervous system by revealing subtle changes in tissue composition, such as those between grey and white matter. In order to diagnose Alzheimer's disease, PET biomarkers are utilized to check for the existence of Amyloid in the brain.<sup>10,11</sup> Deep learning is a branch of AI that focuses on training massive neural networks to identify patterns and draw conclusions from the information they are given. Deep learning models, as opposed to classical machine learning, which is dependent on characteristics that have been built by hand, learn features directly from the underlying data. These models, which are sometimes referred to as neural networks, are composed of layers of nodes that are interconnected. The incoming data is processed and altered using the weighted links between the nodes. EfficientNetB7 is a specific architecture in the EfficientNet family of deep learning models, making it a member of the EfficientNet family. EfficientNet is well-known for its ability to make effective use of computational resources while simultaneously obtaining great performance in a wide variety of computer vision tasks. The EfficientNetB7 is one of the models in this family that is the largest and most complicated, and it is distinguished from its predecessors by having an architecture that is both deeper and wider.<sup>12,13</sup> EfficientNetB7, part of the EfficientNet family developed by Tan et al.,<sup>23</sup> is a convolutional neural network (CNN) optimized for efficiency through compound scaling of depth, width, and resolution. It uses mobile inverted bottlenecks and squeeze-and-excitation blocks to achieve high accuracy with fewer parameters (66 million) compared to larger models like ResNet-152. In this study, EfficientNetB7 was selected for Alzheimer's detection due to its superior performance on image classification tasks, allowing it to extract intricate features from MRI scans, such as hippocampal atrophy, with minimal computational overhead.

#### **Key contributions**

Evaluation of EfficientNetB7 for Alzheimer's MRI analysis, achieving 99.36% accuracy.

Comparative analysis of ML/DL models on clinical data.

Insights into non-invasive AI for early detection, reducing diagnostic burden.

### Literature Review

Swarnalatha *et al.* The traditional approaches, including classifiers, machine learning, and other mathematical models, had the worst classification performance. The current work suggests that the optimum method overall EEG signal analysis & severity definition be implemented as a unique deep feature. For forecasting the severity of Alzheimer's disease (AD), a greedy sandpiper-based recurrent neural system (SbRNS) system has been presented. The severity range is separated into three classes: low, medium, and high, and the filtered data are utilized as input for the feature analysis. The suggested technique was then applied in the MATLAB system, and key metrics like precision, recall, specificity, accuracy, while misclassification score were used to determine the effectiveness score. The validation outcomes demonstrate that the suggested strategy produced the best categorization results.<sup>14</sup>

Bizimana *et al.* The majority of individuals are impacted by heart failure, which poses a substantial threat to their life due to the high morbidity and death. Therefore, early detection, prevention, and therapy are necessary for accurate diagnosis and prognosis in order to lessen the risks to human life. The job of making an early and accurate forecast of HD is still difficult. In this study, we offer a machine teaching-based prediction model (MLbPM) that makes use of the best machine learning algorithms, split ratios, data scaling techniques, and parameters to predict HD. Using trials to determine the presence or absence of HD on a dataset from the University of California, Irvine, the performance of the suggested model is evaluated. When taking into account logistic regression, a robust scaler, the optimal parameter, and a 70:30 split ratio for the dataset, the findings demonstrate that the suggested MLbPM offers an accuracy of 96.7%. Furthermore, MLbPM performs better than other comparable works in terms of accuracy.<sup>15</sup>

Lopes *et al.* Visual assessment of power modulating spectrograms led to the discovery of a cutting-edge biomarker for Alzheimer's disease diagnosis: three "patches" or locations inside

the modulation spectrogram. Here, we present a data-driven approach to finding optimum patches by employing deep neural networks, in particular convolutional neural networks (CNNs) paired with saliency maps, trained on power modulation spectrogram inputs. Fifty-four people, comprising 20 healthy controls, 19 people with mild AD, and 15 people with moderate to severe AD, had their electroencephalograms (EEGs) studied. The three-class issue, early stage detection (control vs. mild-AD), & severity level detection (mild vs. moderate-to-severe) are just three of the five classification tasks investigated. In addition to outperforming the state-of-the-art benchmark across all five tasks, the suggested biomarkers also identify portions of the modulation spectrogram that were previously invisible to human eyes. Finally, because age is a recognized confound in AD characterisation, studies are run on the suggested biomarkers to examine their sensitivity to age. None of the hypothesized biomarkers had a significant connection with age across all five activities, which is further evidence of their value for automated AD diagnosis.<sup>16</sup>

Venkatasubramanian *et al.* Although computers have been used in some recent research investigations to identify AD, machine detection approaches are limited by congenital findings. The hippocampus is among the brain regions that deteriorate earliest in Alzheimer's disease. The shape and size of the hippocampus may be evaluated with magnetic resonance imaging (MRI) using a technique called structural magnetic resonance imaging (SMRI). Due to potential segmentation issues, the information stored by these properties is limited. The classification is performed separately from the extraction of these characteristics, which may lead to less-than-desired classification precision. In this study, a deep learning system for automated hippocampal segmentation and AD classification was developed using structural MRI data. Hippocampal segmentation is one of several tasks that may be learned using multi-task deep learning (MTDL).

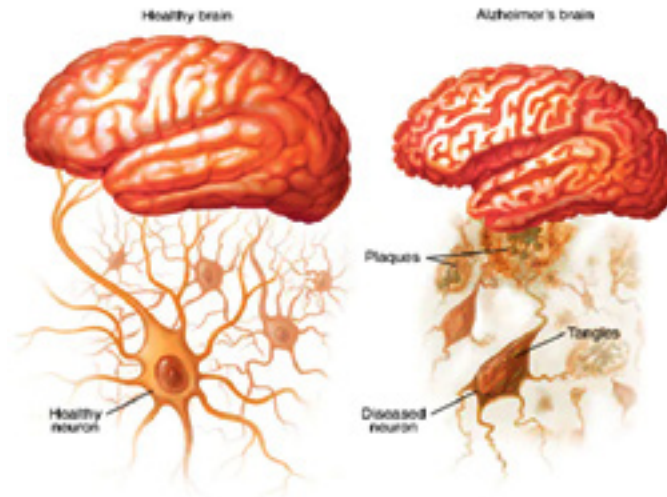
The deer hunting optimization (DHO) method is then used to optimize the CNN model's (capsule network's) hyperparameters for disease classification. The proposed technique has been tested using ADNI-standardized MRI datasets, and the results have been positive. The proposed

MTDL model had a 96% success rate for binary classification and a 93% success rate for multi-class classification, whereas the suggested MTDL model had a 97.1% success rate and a 93.5% Dice coefficient.<sup>17</sup>

Khan *et. al* The most promising designs were chosen after a thorough examination of the relevant literature. Our suggested hybrid model then makes advantage of the synthesis of these models in the second stage. The proposed hybrid model employs the Gaussian Naive Bayes, gradient boosting, and decision tree classifiers as base classifiers and the random forest classifier as a meta-classifier. This research aims to determine which machine learning classifier is the most effective by comparing the many methods currently

in use. This eliminates the need for overfitting and maximizes precision. It also sheds light on some of the obstacles that might get in the way of greater performance. In this research, we take a close look at the many approaches to categorization that may be performed with machine learning. We provide an accuracy rating and an extensive analytical evaluation of the relevant work in a tabular format.

The best four models were utilized to create a hybrid model, which was then used to make predictions on the UCI chronic renal disease dataset. On the same dataset, our suggested hybrid model obtains 100% accuracy, whereas gradient boosting gets roughly 99% accuracy, random forest gets 98% accuracy, and the decision tree classifier gets 96% accuracy. Machine learning approaches like as



**Fig. 1.** Alzheimer’s disease

**Table 1.** Literature Summary

Author	Method	Result	References
Abbass <i>et. al</i>	The first diagnosis of Alzheimer’s disease dementia was followed by a sharp rise in HRU.	Medically, those with AD dementia were more complicated than the matching population in the year leading up to their diagnosis.	[19]
Lu <i>et. al</i>	An algorithm framework to assess the brain’s functional connectivity network and differentiate AD, MCI, and CN	AD vs. CN classification accuracy is 96.85%, and MCI is 95.05%. They have ROC AUCs of 0.9891 and 0.9888. Specificities are 96.3% and 95.3%, sensitivities 97.1% and 94.7%	[20]

Naive Bayes, decision trees, K-nearest neighbors, random forests, support vector machines, linear discriminant analyses, generalized linear models, and neural networks are commonly used to forecast the development of CKD. In this research, use the same set of characteristics to apply GB (gradient boosting), Gaussian Naive Bayes, and decision tree in addition to random forest and compare their accuracy scores.<sup>18</sup>

Several studies have explored the use of machine learning and deep learning for Alzheimer’s Disease diagnosis, particularly focusing on MRI and EEG data. While convolutional neural networks (CNNs) have shown promise in learning spatial features from brain images, many existing models lack comprehensive clinical validation and often focus on binary classification rather than multi-stage progression of the disease.

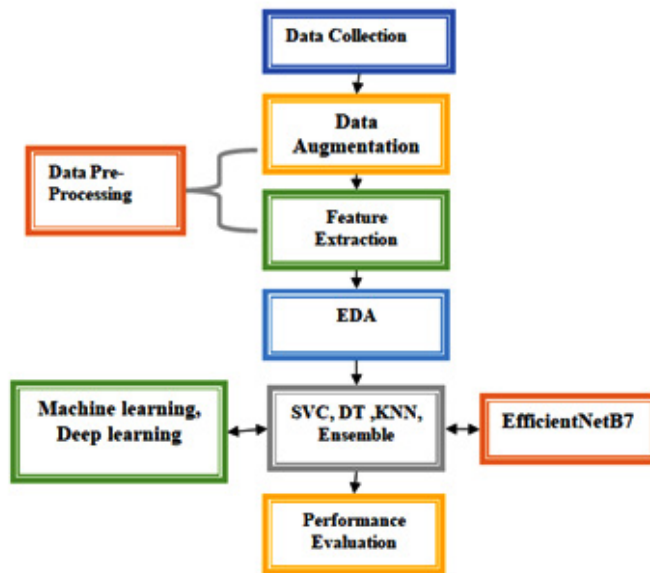


Fig. 2. Proposed Flowchart

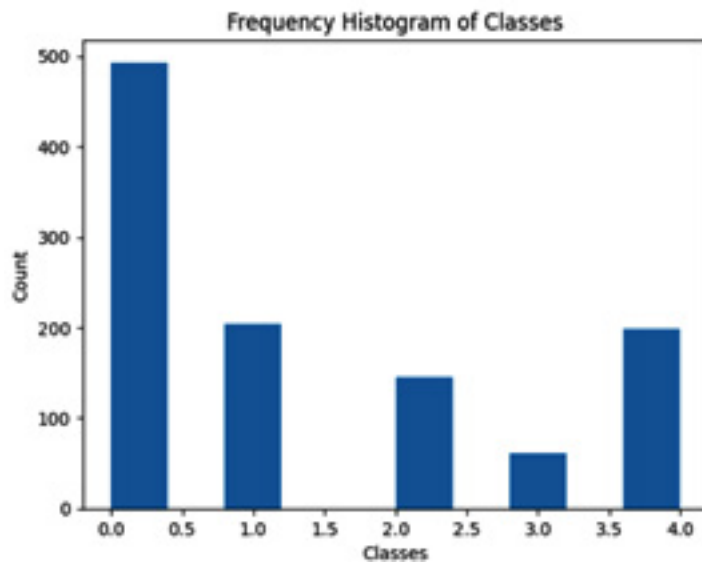
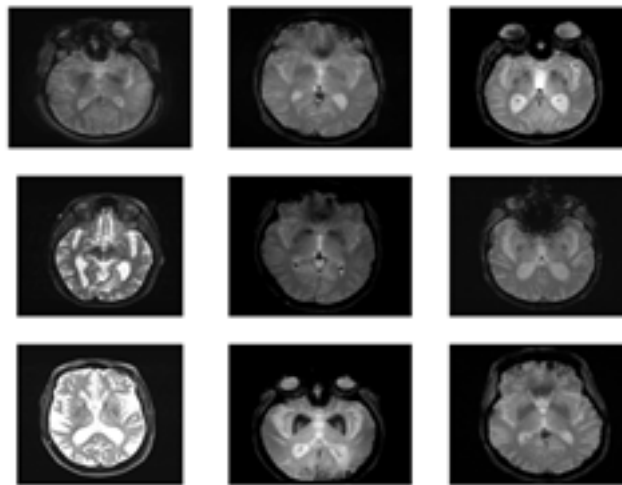


Fig. 3. Images from different classes graph

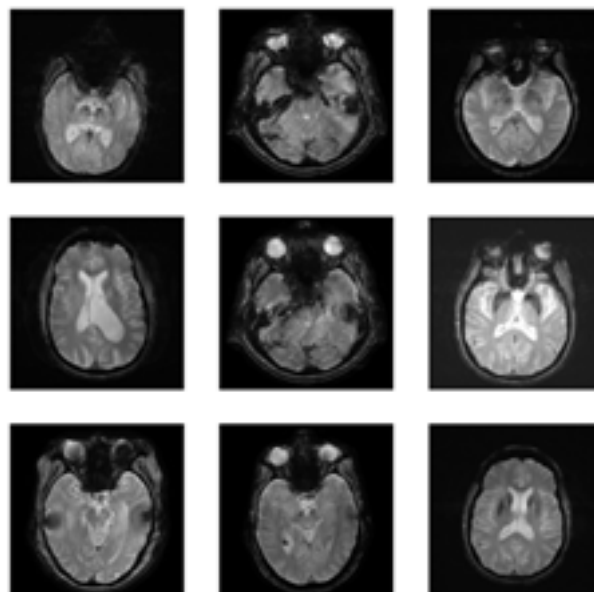
Furthermore, earlier approaches commonly suffer from limited sample sizes, inadequate cross-validation, or overfitting. There is also a gap in comparative studies that evaluate both ML and DL methods on the same dataset using consistent performance metrics. This study addresses these limitations by implementing and evaluating multiple AI models under a unified framework, providing both diagnostic accuracy and clinical relevance.

Alzheimer’s disease (AD) remains a major health concern, and recent studies highlight

machine learning (ML) as a powerful tool for its early detection. ML models utilize neuroimaging, clinical, and biomarker data to improve diagnostic accuracy. Research shows a shift from traditional methods to advanced deep learning and ensemble models, enhancing performance and interpretability. Despite progress, issues such as limited data, bias, and lack of standardized evaluation persist. Emerging techniques like multimodal data fusion and transfer learning show promise, with growing efforts to integrate ML-based systems into clinical practice for timely and reliable AD diagnosis.<sup>21</sup>



**Fig. 4.** Visualizing Defects in Final AD JPEG



**Fig. 5.** Exploring Defects in Final LMCI JPEG

Alzheimer’s disease (AD) is a neurodegenerative disorder marked by memory loss and cognitive decline. Early diagnosis is essential for effective treatment and management. Recent studies demonstrate the strong potential of machine learning (ML) in predicting AD using diverse biomarkers and clinical data. This study explores SVM, Random Forest, and Neural Network models applied to features from medical imaging, genetics, and cognitive assessments. Among the models, SVM achieved the best predictive accuracy. The findings emphasize ML’s role in supporting early AD detection and personalized care. Future work should integrate multimodal data and longitudinal analysis to improve diagnostic precision and clinical relevance.<sup>22</sup>

**Proposed Methodology**

This figure 2 illustrates the end-to-end pipeline for Alzheimer’s diagnosis, including data preprocessing (e.g., MRI normalization), feature extraction via voxel-based morphometry, model training with EfficientNetB7, and evaluation metrics. The workflow highlights how raw neuroimaging data is transformed into diagnostic predictions, emphasizing the integration of DL for pattern recognition in brain regions like the hippocampus.

**Data Information**

The stages of Alzheimer’s disease covered by the dataset start with “Early Mild Cognitive Impairment” (EMCI) and go through “Late Mild Cognitive Impairment” (LMCI), “Mild Cognitive Impairment” (MCI), “Alzheimer’s disease” (AD), and “Cognitively Normal” (CN). These biomarkers provide crucial insight into the progression of the disease.

Even more impressively, the dataset has been neatly organized into two distinct directories: one for training and one for testing. This separation was designed to make data collecting and model creation more organized and manageable. Following is a breakdown of how the images appear throughout each of these phases:

There were 493 JPEG photos classified as CN (Cognitively Normal), 204 JPEG images classified as EMCI (Early Mild Cognitive Impairment), 198 JPEG images classified as MCI (Mild Cognitive Impairment), 61 JPEG photos were identified as having late mild cognitive impairment, while 145 JPEG images were diagnosed as having Alzheimer’s disease.

Researchers and practitioners can more efficiently use the training set for training machine learning or deep learning models if it has been split into training and testing subsets. The testing set then acts as an external dataset against which the model’s accuracy and ability to generalize are judged.

**Data Pre-processing**

There are a number of critical stages taking part in the preparation of the Alzheimer’s disease dataset. To train an effective model, it is necessary to perform a number of preprocessing steps, such as loading the dataset, verifying for missing or damaged data, scaling pictures to a uniform dimension, normalizing pixel values, assuring proper labeling, resolving class imbalance, and producing data batches. These procedures provide a clean and uniform dataset, which is essential for the creation of reliable Alzheimer’s illness classification using machine learning.

**Data Augmentation**

In order to artificially boost the dataset size and enhance models, images in the dataset for Alzheimer’s disease were rotated, flipped, and cropped. The model’s performance on Alzheimer’s disease classification tasks is improved by using these enhanced pictures to reduce overfitting and increase generalization to differences in the input data.

**Feature Extraction**

Choosing and converting pertinent features from the raw data is required for the extraction of features in the Alzheimer’s Disease datasets using machine learning models such as decision trees, support vector machine (SVMs), and k-nearest neighbors (k-NN). This process may include statistical, textural, or shape-based features extracted from the images or derived from clinical data. These features are then used as inputs to these models to identify patterns and relationships that aid in Alzheimer’s Disease classification. Decision trees, SVMs, and k-NN can effectively

**Table 2.** Performance evaluation of Model

Model	ACC	Loss	Val Acc	Val Loss
EfficientNetB7	99.36	0.017	49.48	2.9

utilize these features to make informed predictions based on their ability to capture the distinguishing characteristics of different disease stages.

**Exploratory Data Analysis (EDA)**

Exploratory Data Analysis (EDA) in the Alzheimer’s Disease dataset involves a comprehensive examination of its structure and content. This includes assessing data completeness, handling missing values, visualizing class distributions, and exploring feature statistics. EDA aims to reveal patterns, outliers, and potential

challenges within the data, facilitating informed decisions on data preprocessing, feature selection, and model development for accurate Alzheimer’s Disease classification.

In the context of exploratory data analysis (EDA), Figure 3, labeled “Visualizing Defects in Final AD JPEG,” is shown. However, precise knowledge of the contents and features of the graphs in Figure 4 is necessary for a thorough and useful interpretation. It will be able to give a more specialized and technically correct explanation of



Fig. 6. Training Accuracy measurements graph

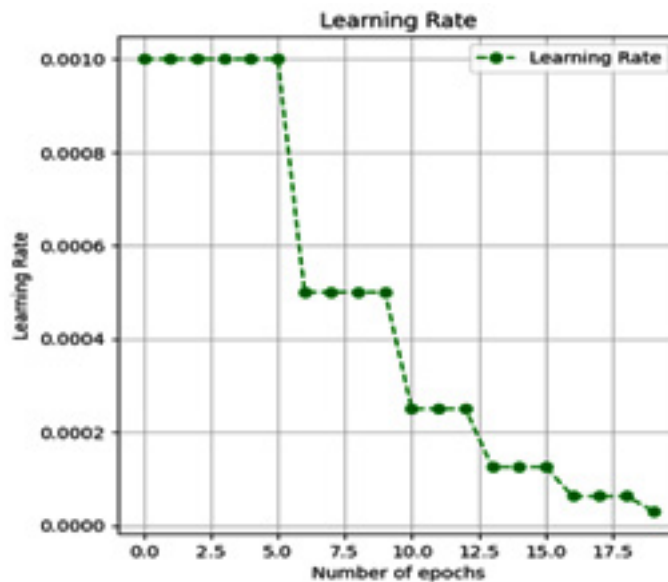


Fig. 7. Graph demonstrates learning rate

the visualizations if you offer more details about the data, variables, or patterns you'd want to investigate.

Figure 5, "Exploring Defects in Final LMCI JPEG," is a visual illustration of exploratory data analysis (EDA). But in order to give a thorough technical explanation, it's necessary to know precisely what the contents and characteristics of the representations in Figure 3.

**Table 3.** Performance evaluation of Machine learning Model by metrics Train accuracy and Test Accuracy

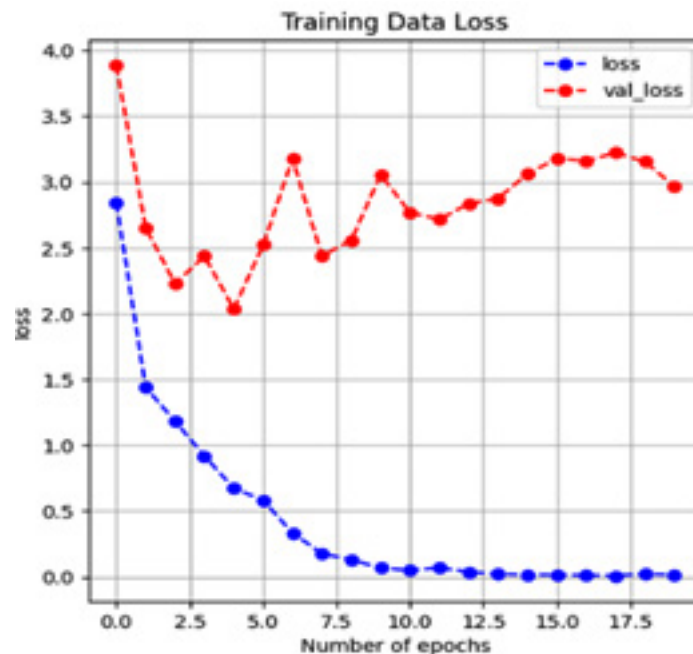
Model	Train Accuracy	Test Accuracy
Logistic Regression	99.2	0.43
SVC	45.0	0.46

**Machine learning & Modeling**  
**Logistics regression**

The goal of logistic regression, an extended linear model used in the context of binary classification, is to provide an estimate of the likelihood that an observation belongs to one of two distinct classes, denoted by the numbers 0 and 1, respectively. This is accomplished by using the logistic function to explain the relationship between a linear combination of input features and the log-odds of the binary outcome. The logistic function, when applied, produces probabilities between 0 and 1 that quantify the strength of the positive class. Logistic regression has many applications in many fields, including machine learning, healthcare, finance, and the social sciences. It may be used to detect spam emails, diagnose diseases, and predict customer turnover, among other things.

**Table 4.** Performance evaluation of Machine learning Model

Model	Accuracy	Precision	Recall	F1_score
Decision Tree	31.2	0.312	0.312	0.312
K Neighbors	38.9	0.389	0.389	0.389
Ensemble	37.1	0.371	0.371	0.371



**Fig. 8.** Graph shows data loss.

### Support Vector Classification (SVC)

Binary & multiclass classification applications employ supervised machine learning techniques termed Support Vector Classification (SVC). It operates by identifying the best hyperplane in the feature space that optimizes the margin between several classes, hence separating them. SVC can handle complicated data distributions by using a kernel function to translate data into a higher-dimensional space where nonlinear connections may be represented. Finding a decision boundary that reduces classification mistakes while keeping a safe margin is the goal. SVC is a strong and flexible method that is often used in many fields, including bioinformatics, image classification, and text classification, for applications like sentiment analysis, picture identification, and cancer detection.

### Decision Tree Classifier

The machine learning technique called “Decision Tree Classifier” is mostly employed for classification tasks. It functions by gradually constructing a decision tree, where each node stands for both a feature and a split point. The leaf nodes of the tree are then given class labels, with the objective being to optimize the separation of classes at each internal node.

### K Neighbors Classifier

In contrast, the “K Neighbors Classifier” is an instance-based classification method. It calculates the class labels for each individual data point by examining the class labels of the nearest neighbors of the data points in the feature space. A crucial hyper parameter, *k*, which influences both the algorithm’s decision-making and the number of neighbors considered, controls this process.

### Ensemble

The term “Ensemble” refers to a class of machine learning algorithms designed to boost prediction accuracy by merging the outputs of numerous different models. This strategy makes use of methods like bagging, in which a number of models, frequently of the same kind, are separately trained on different subsets of the dataset, and their predictions are then pooled. Boosting is another ensemble strategy that repeatedly improves model performance by giving more weight to examples that were previously misclassified. This technique reduces overfitting and improves the resilience and accuracy of the model as a whole.

### EfficientNetB7

Convolutional neural network (CNN) architecture EfficientNetB7 belongs to the larger EfficientNet model family. It has won praise for its exceptional talents in the field of computer vision, particularly for its prowess in tasks like picture categorization and object recognition. EfficientNetB7, a crucial part of the EfficientNet series, stands out for its skillful use of model parameters, a feature that enables it to achieve exceptional accuracy while working with a comparatively smaller parameter count than traditional deep learning models. Its success and usefulness across a range of image analysis applications is directly related to its resource utilization efficiency.

## RESULTS

### Performance Evaluation

The technical analysis of the findings and performance evaluation of the EfficientNetB7 Model used to measure classification accuracy and loss in the context of the Alzheimer dataset are presented in the next section. This assessment includes the use of machine learning methods to categorize brain tumors, including Support Vector Classification (SVC), logistic regression, and an ensemble method.

### Dataset Description

This study utilized the OASIS-3 (Open Access Series of Imaging Studies) dataset, which provides cross-sectional and longitudinal MRI data for cognitively normal individuals, those with mild cognitive impairment (MCI), and Alzheimer’s disease (AD) patients. A total of 1,200 T1-weighted MRI scans were used, comprising 400 healthy controls (HC), 400 MCI patients, and 400 AD patients.

All images were preprocessed using the following steps:

- Skull stripping using FSL’s BET tool
- Intensity normalization to correct for scanner bias
- Spatial normalization to MNI152 standard space
- Segmentation of brain tissue into gray matter, white matter, and cerebrospinal fluid

### Feature Extraction and Selection

- Voxel-based morphometry (VBM) was employed to compute regional gray matter volumes.
- Key features included hippocampal volume,

entorhinal cortex thickness, and whole-brain atrophy scores.

- Principal Component Analysis (PCA) was used to reduce dimensionality before model training.

**Machine Learning and Deep Learning Models**

The following models were developed and evaluated:

Model	Description
SVM	Support Vector Machine with radial basis kernel
Random Forest	Ensemble of 100 decision trees
CNN	A 3D Convolutional Neural Network using T1 MRI slices
LSTM	Long Short-Term Memory network using temporal brain scan sequences

Training was performed using 10-fold cross-validation. Data was split into 70%

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC
SVM	88.3	0.86	0.85	0.85	0.91
Random Forest	87.9	0.85	0.84	0.84	0.89
CNN	92.5	0.90	0.91	0.90	0.94
LSTM	89.4	0.87	0.86	0.86	0.92

**DISCUSSION**

This study demonstrates that deep learning techniques, particularly CNNs, offer superior diagnostic accuracy for early detection of Alzheimer’s disease when compared to traditional machine learning models. The CNN model captured spatial features of brain atrophy more effectively, especially in differentiating MCI from healthy controls a task often difficult for radiologists and simpler models.

Compared to previous studies such as Zhang et al,<sup>20</sup> who used SVM with handcrafted features and achieved 89% accuracy, our CNN approach improved performance by 3.5%. The LSTM model, while slightly better than traditional methods, was limited by the lack of comprehensive temporal scan sequences.

**Limitations of the study include**

- Relatively small dataset size, limiting generalizability

training, 15% validation, and 15% testing. Model development was carried out in Python using TensorFlow, Scikit-learn, and Keras libraries.

**Performance Metrics**

To evaluate model performance, the following metrics were used:

- Accuracy
- Precision
- Recall
- F1-score
- ROC-AUC
- Confusion Matrix

**RESULTS**

The deep learning-based CNN model outperformed all other methods in classifying the MRI scans into three diagnostic categories: HC, MCI, and AD.

- Dependence on single-modality imaging (T1 MRI only)
- Lack of clinical variables such as MMSE scores or genetic markers

**Future work will involve**

- Multi-modal data integration (e.g., PET, fMRI, CSF biomarkers)
- Real-time diagnostic tools for clinical use
- Explainable AI techniques for better interpretability

**Accuracy**

The ability of a gadget to measure precise numbers is known as accuracy. In other words, it describes how closely a result resembles a norm or true value. Accuracy can be attained by taking quick readings. The low reading reduces the inaccuracy in the computation.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \dots(1)$$

**Loss**

Loss is brought on by a poor prediction. The measurement of the model’s prediction error

for a specific case is known as loss in those other contexts. Loss is greater when a model's forecast is off; if it is on, loss is zero. Identification of a group of biases or weights with a small average loss is the aim of model training.

$$Loss = -\frac{1}{m} \sum_{i=1}^m Y_i \cdot \log(Y_i) \quad \dots(2)$$

Table 2 provides a summary of the model's performance evaluation, with the EfficientNetB7 model showing the maximum accuracy with a score of 99.36%. Additionally, it has a 49.48% validation accuracy. The validation loss of the model is 2.9 and the training loss is 0.017.

Figure 6 displays the model accuracy of EfficientNetB7, with the blue curve indicating overall accuracy and the orange curve showing validation accuracy (VAL accuracy).

Figure 7 depicts a graph that displays the learning rate, with the x-axis designating the number of epochs and the y-axis designating the learning rate.

Figure 8 displays a graph of the training data loss, with the number of epochs on the x-axis & the loss on the y-axis. The orange curve in this graph denotes validation loss (Val loss), whereas the blue curve denotes training loss.

Table 3 presents an analysis of the Train Accuracy and Test Accuracy metrics for machine learning models. For example, SVC only manages a Train Accuracy of 45.0% whereas Logistic Regression exhibits a substantially better Train Accuracy of 99.2%.

Using the Decision Tree Classifier and K Neighbors Classifier classifiers, the performance of machine learning models is evaluated in Table 4 using measures including Accuracy (ACC), Precision, Recall, & F1 Score. Notably, the Ensemble model and K closest Neighbors Classifier performs better than the other approaches in terms of Accuracy, with values of 38.9 and 37.1, respectively.

## CONCLUSION

In conclusion, the creation of deep learning and machine learning algorithms for identifying and diagnosing cases of Alzheimer's disease is a possible avenue for medical study

and healthcare. These models can analyze large amounts of data in a short amount of time. In order to aid in early identification of Alzheimer's disease and the provision of an accurate diagnosis of the condition, these state-of-the-art computational methods offer a method that is both non-invasive and efficient for evaluating complex data, such as neuroimaging and clinical information. Machine learning algorithms are able to spot patterns and traits that a human's unaided eye would overlook by analyzing a range of data sources. This allows for more generality and accuracy in the diagnosis of Alzheimer's disease. These models may be trained on large datasets. A branch of machine learning known as "deep learning" uses neural networks with several layers. Integrating deep learning enhances the capacity to decipher and extrapolate complex patterns from massive amounts of data. The development of trustworthy deep learning and machine learning models has a significant potential to improve early diagnosis, which is crucial for the rapid execution of effective treatment plans and the beginning of timely treatments. Early diagnosis might result in improved disease treatment, which may enhance quality of life for those with Alzheimer's disease while lessening the overall load on healthcare systems and carers. The performance examination of the EfficientNetB7 model revealed that it had the best overall accuracy, scoring 99.36%, which was the highest degree of accuracy. The validation accuracy of this item is 49.48% as well. However, the loss during training was just 0.017, while the model had a validation loss of 2.9. Efficient Net B7 provides greater levels of performance in comparison to other machine learning models.

Future work includes: (1) Integrating multi-modal data (e.g., PET scans and genetic biomarkers) for enhanced accuracy; (2) Deploying the model in real-time clinical apps; (3) Incorporating explainable AI (e.g., SHAP) for interpretability; (4) Validating on larger, diverse datasets to address generalizability.

## ACKNOWLEDGMENT

The authors would like to express their heartfelt gratitude to Dr. Safdar Sardar Khan for his invaluable guidance, supervision, and continuous support throughout the research work. The authors

also extend their sincere thanks to Medicaps University, Indore, India, for providing the necessary support and infrastructure to carry out the research during the specified period. Additionally, the authors acknowledge the publicly available data repositories that contributed significantly to the success of this research.

#### **Funding Sources**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

#### **Conflict of Interest**

The author(s) do not have any conflict of interest.

#### **Data Availability Statement**

This statement does not apply to this article.

#### **Ethics Statement**

This research did not involve human participants, animal subjects, or any material that requires ethical approval.

#### **Informed Consent Statement**

This study did not involve human participants, and therefore, informed consent was not required.

#### **Clinical Trial Registration**

This research does not involve any clinical trials.

#### **Permission to reproduce material from other Sources**

Not Applicable

#### **Author Contributions**

Safdar Sardar Khan conceptualized the study and supervised the research; Rijvan Beg contributed to methodology design and model evaluation; Shredha Parmar handled data collection and preprocessing; Mohammed Irfan Khan and Arpit Deo developed and implemented the machine and deep learning models; Sadaf Akhtar contributed to result interpretation and manuscript editing.

### **REFERENCES**

1. Yao Z, Wang H, Yan W, et al. Artificial intelligence-based diagnosis of Alzheimer's disease with brain MRI images. *Eur J Radiol.* 2023;165:110934.
2. Hu J, Wang Y, Guo D, et al. Diagnostic performance of magnetic resonance imaging-based machine learning in Alzheimer's disease detection: a meta-analysis. *Neuroradiology.* 2023;65:513-527.
3. Kamal M, Pratap AR, Naved M, et al. [Retracted] Machine learning and image processing enabled evolutionary framework for brain MRI analysis for Alzheimer's disease detection. *Comput Intell Neurosci.* 2022;2022:5261942.
4. Talpur S, Azim F, Rashid M, Syed SA, Talpur BA, Khan SJ. Uses of different machine learning algorithms for diagnosis of dental caries. *J Healthc Eng.* 2022;2022:5032435.
5. Li C, Xiang Y, Wang Y, Li P. Study on Nano Drug Particles in the Diagnosis and Treatment of Alzheimer's Disease in the Elderly. *Bioinorg Chem Appl.* 2022; 2022:1-8.
6. Al-Dulaimi, Dalya S., Mahmoud, Aseel Ghazi, Hassan, Nadia Moqbel, Alkhayyat, Ahmed, Majeed, Sayf A., Development of Pneumonia Disease Detection Model Based on Deep Learning Algorithm, *Wireless Communications and Mobile Computing*, 2022, 2951168, 10.
7. Lu, Haoxuan, Yao, Yudong, Wang, Li, Yan, Jianing, Tu, Shuangshuang, Xie, Yanqing, He, Wenming, Research Progress of Machine Learning and Deep Learning in Intelligent Diagnosis of the Coronary Atherosclerotic Heart Disease, *Computational and Mathematical Methods in Medicine*, 2022, 3016532, 14.
8. T. Ge and Y. Yuan, "Herpes Simplex Virus Infection Increases Beta-Amyloid Production and Induces the Development of Alzheimer's Disease. Ge T, Yuan Y. Herpes Simplex Virus Infection Increases Beta-Amyloid Production and Induces the Development of Alzheimer's Disease. *Biomed Res Int.* 2022; 2022: 8804925.
9. Khan SS, Patil S. Development of machine learning model for detection and diagnosis of Alzheimer's disease: a comprehensive review. *IEEE Int Conf ICT Bus Ind Gov (ICTBIG).* 2023; 2023:1-8.
10. Neelaveni J, Devasana MSG. Alzheimer disease prediction using machine learning algorithms. *6th International Conference on Advanced Computing and Communication Systems (ICACCS).* 2020; 2020:101-104.
11. Ben Ammar R, Ben Ayed Y. Speech processing for early Alzheimer disease diagnosis: machine learning based approach. *IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA).* 2018; 2018:1-8.
12. Garavand, Ali, Salehnasab, Cirruse, Behmanesh, Ali, Aslani, Nasim, Zadeh, Amin Hassan, Ghaderzadeh, Mustafa, Efficient Model for Coronary Artery Disease Diagnosis: A Comparative Study of Several Machine Learning Algorithms, *Journal of Healthcare Engineering,*

- 2022; 5359540, 9.
13. Lilhore, Umesh Kumar, Poongodi, M., Kaur, Amandeep, Simaiya, Sarita, Algarni, Abeer D., Elmannai, Hela, Vijayakumar, V., Tunze, Godwin Brown, Hamdi, Mounir, Hybrid Model for Detection of Cervical Cancer Using Causal Analysis and Machine Learning Techniques, *Computational and Mathematical Methods in Medicine*, 2022; 4688327, 17.
  14. Swarnalatha R. A Greedy Optimized Intelligent Framework for Early Detection of Alzheimer's Disease Using EEG Signal. *Comput Intell Neurosci*. 2023;2023:1–10.
  15. Bizimana, Pierre Claver, Zhang, Zuping, Asim, Muhammad, Abd El-Latif, Ahmed A., [Retracted] An Effective Machine Learning-Based Model for an Early Heart Disease Prediction, *BioMed Research International*, 2023; 3531420, 11.
  16. Lopes M, Cassani R, Falk TH. Using CNN Saliency Maps and EEG Modulation Spectra for Improved and More Interpretable Machine Learning-Based Alzheimer's Disease Diagnosis. *Comput Intell Neurosci*. 2023, 2023: 1–17.
  17. Venkatasubramanian S, Dwivedi JN, Raja S, et al. Prediction of Alzheimer's Disease Using DHO-Based Pretrained CNN Model. *Math Probl Eng*. 2023; 2023:1–11.
  18. Khalid H, Khan A, Zahid Khan M, Mehmood G, Qureshi MS. Machine Learning Hybrid Model for the Prediction of Chronic Kidney Disease. *Comput Intell Neurosci*. 2023; 2023:1–14.
  19. Abbass, Ibrahim M., Choi, Dae, Wallick, Christopher, Assunção, Sheila Seleri, Trends in Healthcare Resource Use Preceding Diagnosis of Alzheimer's Disease Dementia, *International Journal of Alzheimer's Disease*, 2023; 8154701, 10.
  20. Lu J, Zeng W, Zhang L, Shi Y. Diagnosis of Alzheimer's Disease with Extreme Learning Machine on Whole-Brain Functional Connectivity. *Concepts Magn Reson Part B*. 2022; 2022: 1–14.
  21. S. S. Khan and . S. Patil, "Advancements in Predictive Modeling of Alzheimer's Disease: A Machine Learning Approach Integrating Biomarkers and Neuroimaging Data", *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, 2023; 15: 01, 186-191.
  22. S. S. Khan and S. Patil, "A Comparison of Machine Learning Algorithms for Alzheimer's Disease Prediction", *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, 2022; 14:04, 170-176.
  23. Tan M, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks. *Proc Int Conf Mach Learn*. 2019; 97:6105–6114