International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(15s) 2026 pp. 103-119
https://doi.org/10.70917/ijcisim-2026-4351
Article

Multimodal Deep Learning for Early Alzheimer's Disease Prediction Using MRI Texture Features and Gut Microbiome Data

Anil Chakravarthy[1*], Sasmita Kumari Nayak[2], B S Panda[3]

[1]Department of Computer Science & Engineering, Centurion University of Technology and Management Odisha, India.
Email id: anilchakravarthy2015@gmail.com
[2]Department of Computer Science & Engineering, Centurion University of Technology and Management Odisha, India.
Email id: nayaksasmita484@gmail.com
[3]Department of Computer Science & Engineering, Raghu Engineering College, Visakhapatnam, India.
Email id: bspanda.cse@gmail.com

Abstract: Early diagnosis of Alzheimer Disease (AD) has become one of the most significant problems in clinical neuroscience since structural and biological changes may occur decades before a symptom is visible to the naked eye. In this paper, a new multimodal Deep Learning (DL) model is suggested, which incorporates MRI texture features based on Gabor filters and gut microbiome data to better predict of Alzheimer disease in early. The structural brain changes are recorded by the use of Magnetic Resonance Imaging (MRI) and the Gabor filters may be applied to extract the discriminative texture features that have the capability of enhancing the faint profiles which are associated with the initial neurodegeneration. Simultaneously, microbial composition and diversity data, which are the gut microbiome, is added as a new biomarker associated with the gut-brain axis and cognitive impairment. The proposed model combines a hybrid DL architecture that involves the use of convolutional neural networks (CNNs) to handle Gabor-enhanced MRI data and fully connected layers to handle microbiome features data. Multimodal fusion approach is applied so that the heterogeneous sources of data can be properly combined so that the model could acquire complementary representations. The framework is experimented on the categorization at the initial level, between cognitively normal, mild cognitive impairment (MCI) and the Alzheimer disease. The evidence of the experiment proves that imaging and microbiome data integration are much more effective than unimodal in predicting the results. In addition, the sensitivity of the features to the primary structural changes of brain tissues is enhanced by Gabor filters. The mentioned study identifies the opportunities of the combination of neuroimaging and biological data with a DL to assist in the early diagnosis of the Alzheimer disease as a potential course in the precises medicine and preventive care.
Keywords: Alzheimer's Disease Prediction, Multimodal Deep Learning, MRI Texture Analysis (Gabor Filter),Gut Microbiome, Early Detection / Mild Cognitive Impairment (MCI).

1. Introduction

AD is a progressive type of neurodegenerative disease, and one of the leading causes of dementia in the world and as such, it impacts significantly on cognitive ability, memory and daily living activities. Due to the increasing number of the aged population in the world, the burden of the Alzheimer disease is bound to increase tremendously, leading to severe challenges in the healthcare systems. Early AD diagnosis, particularly in the MCI is vital in the process of ensuring the disease is mitigated and treated in time [1]. Nevertheless, the conventional methods of diagnosis usually entail clinical manifestations which manifest themselves at late stages and restrict the usefulness of early treatment methods.

Recent advances in neuroimaging have enabled the possibility of identifying structural changes in the brain with reference to Alzheimer disease. MRI has widely been applied to detect atrophy of the brain and other abnormalities especially brain regions such as the hippocampal regions. Nonetheless, slight alterations of brain tissues in their texture at the early stage are frequently hard to record with the help of traditional technique [2]. To overcome

Copyright: © 2026 by the authors

this shortcoming, methods of texture analysis including Gabor filters have been considered to have strength in extracting multi-scale and multi-orientation characteristics to promote the detection of fine structural variations in MRI images.

Simultaneously, new studies denote the involvement of gut–brain axis in neurodegenerative diseases. The gut microbiome, which is a composition of trillions of microorganisms, has been reported to mediate brain activity in immune, metabolic and neural pathways [3]. Changes to the microbial composition have been linked to cognitive impairment and Alzheimer disease, and thus the gut microbiome is a promising non-invasive biomarker to diagnose the disease at an early stage. In spite of this, microbiome data is not used adequately in predictive modeling of AD.

To leverage the complementary property of these modalities, the present paper proposes the introduction of a multimodal deep learning framework, which uses Gabor-enhanced of MRI texture information with gut microbiome information [4]. The suggested method will aid the accuracy of the prediction at an earlier level by introducing structural data with the application of the imaging and biological signatures. The heterogeneous information can be effectively learned with the DL algorithms, CNN and feature fusion strategies and the complex patterns are not always apparent using the traditional methods.

The study will help in the creation of a strong and explainable predictive model of the early identification of Alzheimer [5]. The study has also given a new direction of precision medicine and data-driven healthcare by combining neuroimaging and microbiome analysis, which has the possibility of making early diagnosis and enhancing patient outcomes [6].

2. Literature Survey

Gottipati et al., (2024) in [7] offered a model of brain tumors classification that combines Classical Local Binary Patterns (CLBP) with the CNN to analyze MRI images successfully. CLBP is employed to obtain texture content and other traditional methods such as Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG), which enables texture content to reveal local and global patterns that play a vital role in tumor identification. The proposed algorithm will involve preprocessing of the images, feature extraction using the CLBP and classification using the CNN with a high accuracy rate of 95.6 percent on a publicly available set of images. The results confirm that deep learning-based texture-based feature extraction can improve the diagnostic performance. Other forms of classification such as ANN, Adaptive Differentiating Evolution (AlDE) and Linear Discriminant Analysis (LDA) have also been discussed in the paper. Future recommendations: It is suggested that further studies should be conducted to enhance the feature extraction by using such variants as Dynamic LBP (DLBP) and extended LBP methods to enhance the classification performance and resistance in medical image analysis.

Parvatham and Lakshmana (2025) in [8] proposed study suggests a multimodal deep learning model to identify the Alzheimer disease by means of structural MRI (sMRI) and functional MRI (fMRI) data. The SMRI structural properties are acquired with the help of CNN based models, and the patterns of the brain connectivity are acquired with the help of fMRI. The framework utilizes the state of the art deep learning architectures such as GoogLeNet, DenseNet-121, Graph Neural Network (GNN) and U-Net, which are stacked using the ensemble strategy. The probabilities of prediction of these models are vertically stacked to generate a meta-feature matrix and this is then classified with a Random Forest meta-model. The proposed approach is premised on the structural and functional brain data, which enables the enhancement of the accuracy of classification and generalization. The model was highly precise and scored 95.18 that was above the current state of art which had great chances of identifying early Alzheimer disease and neurodegenerative studies.

Abbas, S et al., (2025) in [9] presented a hybrid deep learning model, composed of GLCM, VGG16, FMO and CNN-LSTM, to enhance the diagnosis of Alzheimer disease with the aid of MRI images. The methodology involves preprocessing that involves normalization and noise removal, feature extraction in terms of texture features through the use of GLCM and spatial features through the use of VGG16. Fisher Mantis Optimization (FMO) is used to select the best features and the chosen features are categorized with the help of CNN-LSTM model in order to utilize both space and time features. The model is tested using the ADNI and MIRIAD data, where it has an excellent performance of 98.63 percent accuracy, high sensitivity, precision, and F1-score. The proposed methodology beats the current models, including CNN + SVM and 3D-CNN + BiLSTM, and shows the superiority of FMO to the other optimization algorithms, and it is resistant to varying hyperparameter configurations.

Karve, A (2025) in [10] proposed a machine learning-driven system of early detection of Alzheimer disease, which makes use of Decision Tree, Random Forest, and Gradient Boosting algorithms. The model is built with the

preprocessed data of the ADNI and OASIS data sets, including the structural MRI data, cognitive tests and biomarkers data. Ensemble techniques, especially Random Forest and Gradient Boosting, have a better performance than the traditional ones, with the maximum classification accuracy of 89 and better precision, recall, and F1-score. Also, the Explainable AI methods (SHAP and LIME) increase the interpretability of the model enabling clinicians to understand the main determinants of predictions. The paper evaluates the significance of integrating machine learning with explainability in early diagnosis and informed clinical decision-making, and recommends future research with longitudinal and multimodal data to achieve higher levels of robustness.

Hechkel W & Helali (2025) in [11] suggested a new computer-aided diagnostic model that can be used to detect and classify the early stages of the AD with the help of the YOLOv11 neural network. The method combines multimodal information of T2-weighted MRI and DTI images provided in the ADNI database that makes it possible to locate and classify AD-related biomarkers simultaneously. Interests areas are identified based on known biomarkers and five stages are segregated to which subjects are classified i.e. cognitively normal, early MCI, late MCI, MCI. The proposed method produces promising performance of 93.6% precision, 91.6% recall and 96.7% mAP50 indicating its usefulness in discovering latent trends of the disease. The paper sheds light on the possibility of the YOLOv11 being used in mixed tasks of detecting and classifying that could become the future promise in the diagnosis of early-stage Alzheimer.

S. Shanto and N. Haque (2024) in [12] introduced a new computer-aided diagnostic method to identify early signs of Alzheimer disease with the help of the YOLOv11 neural network. The process entails integration of multimodal information of T2-weighted MRI and DTI images to concomitantly localize and classify AD-related biomarkers. The areas of interest are set based on the presence of bio-markers and the model classifies subjects into four stages that comprise cognitively normal, early MCI, late MCI and MCI. The method proposed offers excellent performance in terms of high precision, recall and mAP50 and is able to detect subtle patterns of the disease. The study shows the abilities of YOLOv11 in combined detection and classification, which is an effective approach to diagnosing Alzheimer at its early stages.

A.P. Porsteinsson et al., (2021) in [13] suggested a computationally effective CNN-based model to predict the presence of Alzheimer disease in its initial stages and classify it based on T1-weighted 2D MRI images of the ADNI dataset. The model is used to categorize the subjects into cognitively normal, stable mild cognitive impairment, and Alzheimer disease. Despite the low computational complexity, the proposed CNN has a high accuracy of 98.97% that is comparable with majority of the existing methods and indicates the efficiency of lightweight deep learning architecture in terms of accuracy and low cost in Alzheimer detection.

Kim et al., (2025) in [14] highlighted the significance of early and correct diagnosis of the Alzheimer disease to enable proper diagnosis, management, and better quality of life of the patient. It describes the challenges of diagnosing AD at an early stage in clinical practice including lack of time by the clinicians, problems in diagnosis and misunderstanding of symptoms as normal ageing. The article notes the need to improve the diagnostic approaches and interdisciplinary treatment in the clinical profession, starting with primary care. It also provides some practical suggestions and avenues to guide the medical professionals through the diagnostic process in a bid to encourage a more efficient and comprehensive process of the identification of the Alzheimer at its early stages.

Miltiadous, A et al., (2023) in [15] examined how the spectrum of the Alzheimer disease can be classified, that is, subjective cognitive decline (SCD), MCI and AD by examining EEG signals collected under resting-state and task-based conditions. EEGs were captured when the participant was performing a simple cognitive task that included listening to the task, comprehending it, and responding to it, and rest-state measurements. The attention-based LSTM and EEG Conformer DL models were used in binary and multi-class classification. The results obtained indicate that task based EEG is significantly better than resting-state EEG the accuracy of task based is 5-15% better and also the value of AUC is also higher and no significant difference is seen between the two models. The paper recognizes the performance of task-based EEG along with deep learning to enhance the early detection of the Alzheimer disease.

3. Proposed Work

The proposed research will be dedicated to the development of a multimodal deep learning system to predict AD at the early stages with the help of MRI texture features obtained with the assistance of Gabor filters and gut microbiome data [16]. The framework will be such that it will be able to capture the structural abnormalities of the brain and the pattern of microbial imbalance that when combined will be able to enhance the discrimination of CN, MCI and AD subjects.

3.1 Data Collection

The data employed in this paper is derived on Kaggle and is made up of a total of 7200 images of the human brain MRI divided into four separate categories; glioma tumor, meningioma tumor, pituitary tumor and no tumor with 1800 images per category. These classes are typical brain tumor models and normal brain images, and the model is capable of learning the pathological and non-pathological models [17]. The data is grouped into fixed training and testing sets, which guarantee a systematic process of model training and testing. The dataset is also balanced with the classes, which makes it possible to prevent bias during the model training, and also improves the classification of all classes. Figure 1 below is the sample picture that is taken out of the above defined data.

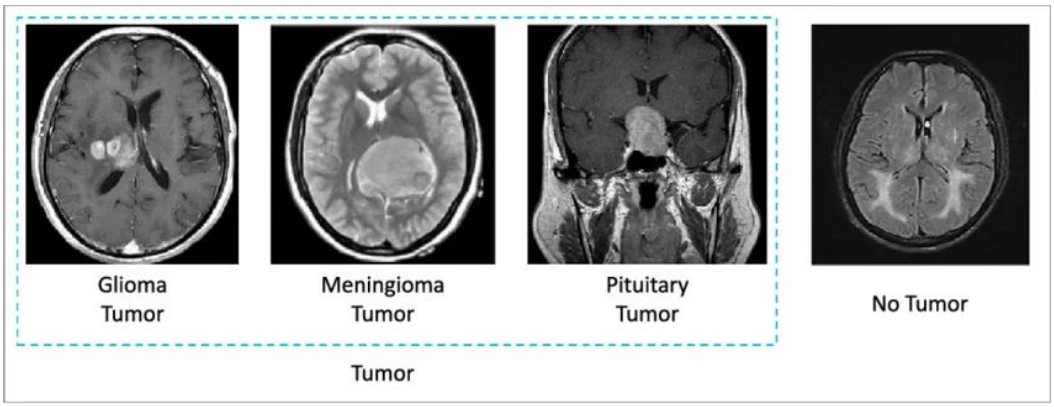

Figure 1: Sample image from dataset

3.2 Pre-processing

To enhance the quality of the MRI analysis, a new adaptive preprocessing pipeline is suggested, which comprises of skull stripping, intensity normalization, noise removal, and region of interest extraction. All steps are aimed at maintaining fine anatomical features and enhance data consistency.

i. Novel Skull Stripping Method

A new Adaptive Boundary-Aware Skull Stripping (ABASS) is put forward to precisely divide brain tissue and non brain tissue like skull, fat and scalp. This technique eliminates the non-brain tissues using intensity thresholding and edge guided refining of the boundaries [18]. The brain mask is created by retaining the most anatomically interrelated part.

$$M(x,y) = \begin{cases} 1, & if\ I(x,y) > \tau\ and\ \left|\left|\nabla I(x,y)\right|\right| < \delta \\ 0, & otherwise \end{cases} \quad (1)$$

Where $M(x,y)$ is the brain mask, τ is the adaptive intensity threshold, and δ is the gradient constraint.

the mask of the brain, τ is the adaptive intensity threshold and δ is the gradient constraint.

The stripped image is:

$$I_{ss}(x,y) = I(x,y).M(x,y) \quad (2)$$

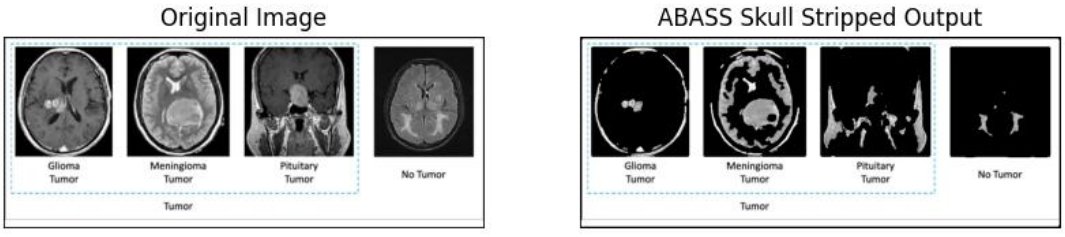

Figure 2: Original Vs Skull stripped image

ii. Intensity Normalization

When skull stripping is done, the brain image is divided into coarse tissue zones Tissue-Adaptive Intensity Normalization (TAIN). This technique balances the intensity of the MRI images by statistical adjustments of the tissue to minimize inter-scan error but maintain contrast.

$$I_{norm}(x,y) = \frac{I_{ss}(x,y)-\mu_t}{\sigma_t} \quad (3)$$

Where, μ_t and σ_t are the standard deviation and the mean of the respective tissue region t.

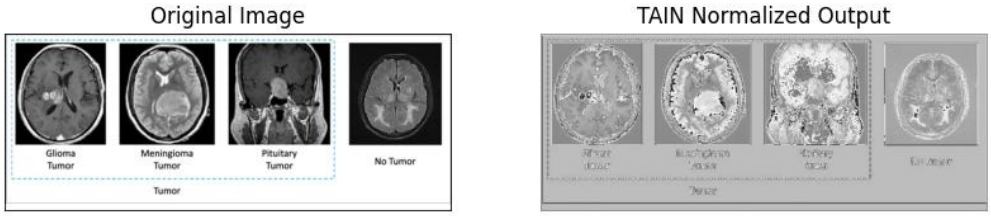

Figure 3: Original vs Normalized image

iii. Noise Removal

To remove noise, a Structure-Preserving Neuro-Adaptive Denoising (SPNAD) is suggested. The noise can be suppressed by conventional smoothing filters and tend to blur fine structural features that are very important in detecting early abnormalities. The technique eliminates noise and preserves fine anatomic structures through local variance-controlled smoothing.

$$I_{den}(x,y) = I_{norm}(x,y) - \alpha(x,y)N(x,y) \quad (4)$$

$N(x,y)$ is the estimated noise, and $\alpha(x,y)$ is an adaptive weight, which is dependent on local texture variance.

$$\alpha(x,y) = \frac{1}{1+Var_{local}(x,y)} \quad (5)$$

This guarantees that there is greater denoising in smooth areas and reduced smoothing in smooth areas.

iv. Region of Interest Extraction

The proposed novel Attention-Based Disease-Sensitive ROI Extraction (ADS-ROI) approach can detect anatomically and clinically relevant brain regions related to the initial neurodegenerative alterations [19]. The suggested approach will automatically detect multiple candidate regions based on their sensitivity to structural abnormalities and not select a single fixed region. This technique determines clinically relevant regions of the brain based on a weighted sum total of texture abnormality and intensity deviation.

$$A_r = w_1 T_r + w_2 D_r \quad (6)$$

with A_r indicating the attention rating of region r, T_r indicating the texture response, D_r indicating the intensity deviation, and w_1, w_2 being weights. The ROI is selected as

$$RoI = \{r | A_r > \lambda \quad (7)$$

where λ is the selection threshold.

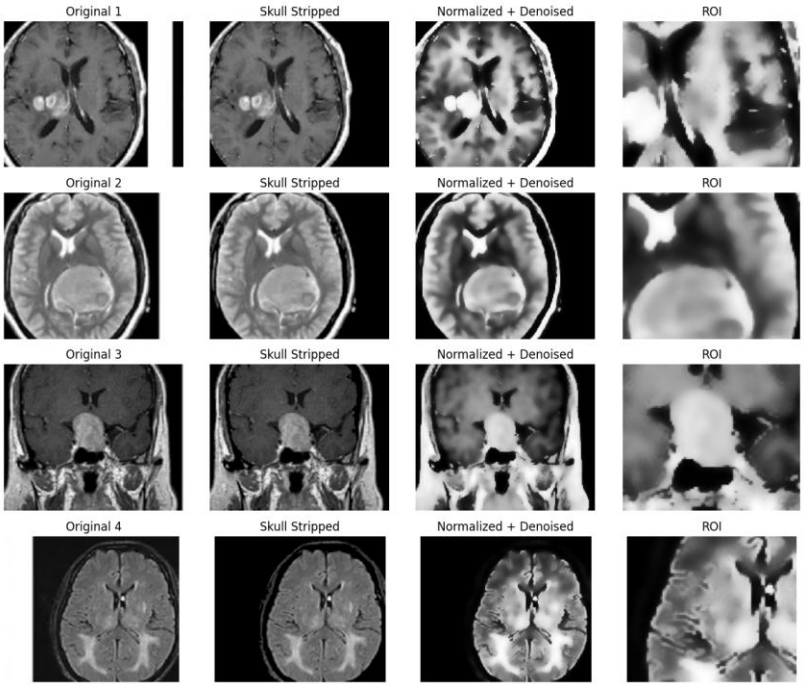

Figure 2: Original Vs Pre-processed image

3.3 Feature Extraction

The extraction of texture features forms an important step in image analysis using MRI as pathological areas usually have minor structural and intensity differences that cannot be adequately represented using raw pixel values only. Gabor filters are used to remove discriminating texture data in preprocessed MRI scans to form a proposed framework. Gabor filtering is more applicable in the analysis of medical images as joint localization is attained in both the spatial and the frequency domain, which facilitates the successful characterization of orientation dependent and scale dependent tissue patterns. This renders it highly applicable in detecting abnormalities of heterogeneity of tumours, tissue abnormality, and microstructural deviation [20].

A two-dimensional (2D) Gabor filter can be said to be a plane wave that is a sinusoid modulated by a Gaussian envelope. The overall shape of a 2D Gabor filter is provided as follows:

$$G(x, y; \lambda, \theta, \varphi, \sigma, \gamma) = exp - \frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \cos 2\pi \frac{x'}{\lambda} + \varphi \quad (8)$$

Where,

$$x' = xcos\theta + ysin\theta \;\; y' = -xsin\,\theta + ycos\theta \quad (9)$$

Where λ is the wavelength of the sinusoidal component, θ is the orientation of the filter, φ is the phase offset and σ is the spread of the Gaussian envelope and γ is the spatial aspect ratio.

These parameters enable the filter to be selective to image structures of a particular frequency and directional characteristics.

Assume that the processed MRI picture is represented as I I(x,y)I(x,y)I(x,y). The multiple scales and orientations are used to build a bank of Gabor filters to represent different textural features. Final image is the result of the convolution between the output of the MRI image and Gabor filter of scale m and orientation n.

$$R_{m,n}(x, y) = I(x, y) * G_{m,n}(x, y) \quad (10)$$

where $*$ denotes the convolution operation. The size of the filter response is then computed as:

$$M_{m,n}(x, y) = |R_{m,n}(x, y)| \quad (11)$$

The magnitude response of an image texture indicates the amount of energy of the image texture and is frequently utilized in feature extraction due to its insensitivity to small phase shifts. Using a filter bank consisting of M scales and N orientations, one will get a set of Gabor response maps:

$$R = \{M_{m,n}(x,y)\}, \quad m = 1,2,\ldots.M, n = 1,2,\ldots..N \tag{12}$$

These response maps code the complementary information of the MRI texture of various resolutions and orientations. A set of statistical texture descriptors are obtained out of each Gabor response map to create the final feature vector. Mean, standard deviation, energy, and entropy are examples of commonly used descriptors.

i. Mean

The average of the response map is calculated as:

$$\mu_{m,n} = \frac{1}{PQ}\sum_{x=1}^{P}\sum_{y=1}^{Q} M_{m,n}(x,y) \tag{13}$$

where P and Q are the dimensions of response map.

ii. Standard Deviation

It quantifies the dispersion of intensity of values of textures. The standard deviation is obtained as:

$$\sigma_{m,n} = \sqrt{\frac{1}{PQ}\sum_{x=1}^{P}\sum_{y=1}^{Q}(M_{m,n}(x,y) - \mu_{m,n})} \tag{14}$$

iii. Energy

The strength of the texture patterns is given as the energy feature which is defined as:

$$E_{m,n} = \sum_{x=1}^{P}\sum_{y=1}^{Q} M_{m,n}^2(x,y) \tag{15}$$

As the entropy, a measure of textural randomness and irregularity, is provided by,

$$H_{m,n} = -\sum_k p_k \log(p_k) \tag{16}$$

where p_k is the chance of the occurrence of intensity level k in Gabor response map histogram.

Algorithm 1: Feature Extraction using Gobar filter

Step 1: Feed Input MRI image

Step 2: Texture feature Extraction using Equation (8)

Step 3: use filter to perform Convolution by using equation (10)

Step 4: To make features robust, use magnitude with equation (11), This is the strength of texture at a pixel.

Step 5: Extraction of statistical feature through equation (13) to (16)

Step 6: Feature Vector Form

All features are combined and represented as:

$F = [\mu, \sigma, E, H]$ at all scales and orientations.

Final feature size can be computed and expressed in the form of

$$D = (no\ of\ scales) * (no\ of\ orientations) * features$$

Step 7: final list of sample feature is extracted.

3.4 Gut microbiome preprocessing

Preprocessing of gut microbiome is an important step in multimodal prediction of Alzheimer's disease because raw microbiome data produced by sequencing (e.g. 16S rRNA or metagenomics) is high-dimensional, sparse,

compositional and noisy. Adequate preprocessing guarantees that before giving the data to machine learning models, biologically relevant microbial patterns are maintained and variability and redundancy is minimized.

Step 1: Microbiome Data Representation The microbiome data is represented in the initial step.

Each microbiome in the gut of the subject is modeled as a feature vector:

$$X^{GM} = [g_1, g_2, \dots g_d] \quad (17)$$

where:

g_j= presence of the j^{th} microbial taxon (species/genus), d total number of taxa.

Step 2: Handling Compositional Data

The data in microbiomes is compositional (i.e., proportions but not absolute counts). It must be normalised, then:

$$g'_j = \frac{g_j}{\sum_{k=1}^{d} g_k} \quad (18)$$

The step eliminates sequence depth bias and normalizes samples.

$$\sum_{j=1}^{d} g'_j = 1 \quad (19)$$

Step 3: Transformation (Variance Stabilization) into logs.

Microbiome data tend to be very skewed. To reduce skewness:

$$g''_j = \log(g'_j + \in) \quad (20)$$

$\in$ is a small constant (e.g., 10−610^{-6}10−6) to avoid log(0).

Step 4: Data Transformation Compositional Data (CLR Transformation)

Microbiome data exists in the simplex space hence Centered Log-Ratio (CLR) transformation is used:

$$CLR(g_j) = \log \frac{g'_j}{\sum_{k=1}^{d} g'_k)^{\wedge}\frac{1}{d}} \quad (21)$$

denominator is the geometric average of all the taxa, This eliminates compositional problems and permits conventional ML models.

Step 5: Dealing with Sparsity or Zero Imputation.

Large datasets of microbiomes have numerous zeros because of rare species. These are processed through pseudo-count addition:

$$g^*_j = g_j + \in \quad (22)$$

Eliminates unspecified logarithmic functions and makes it stronger.

Step 6: Dimensionality Reduction / Feature Selection.

The sizes are very high, and thus only informative microbial features are selected by specifying the following:

$$MI(g_j, y) = \sum_{g_j} \sum_{y} p(g_j, y) log \frac{p(g_j, y)}{p(g_j)p(y)} \quad (23)$$

Based on the above computation choose top K features,

$$F^{GM} = [g_1, g_2, \dots . g_k] \quad (24)$$

Step 7: Scaling and Standardization.

To have equal contribution of features:

$$g'_j = \frac{g_j - \mu_j}{\sigma_j} \quad (25)$$

$\mu_j$is feature and σ_j is standard deviation.

Step 8: Final Microbiome Feature Vector

After pre-processing the final feature vector is computed from $F^{GM} = [g'_1, \ldots g'_k]$. The feature vector obtained is then used as input into the ML model to do additional classification.

3.5 Feature Learning

The multimodal feature learning task entails the extraction and learning of complementary representations of heterogeneous data sources, in this case, MRI texture maps and gut microbiome features, to enhance the early detection of Alzheimer disease. Given that these modalities have different formats (image vs. tabular), dedicated deep learning models are applied to each, and they are integrated into a single framework.

3.5.1 CNN-Based Feature Learning for MRI Texture Maps

The MRI image is first filtered with Gabor filters after which it is converted into texture-enhanced feature maps that are taken as input to a CNN to learn hierarchical features. Multimodal feature learning Multimodal feature learning is developed to learn to represent the heterogeneous data, i.e., MRI texture maps and gut microbiome features, together to enhance the early prediction of Alzheimer disease. These modalities are very different in their structure, i.e., image-based and tabular, and therefore different branches of deep learning are applied to extract modality-specific features that are subsequently combined into a single representation.

3.5.1.1 CNN-Based Deep Feature Learning for MRI Texture Maps

The MRI images are preprocessed and Gabor filtered after which they are converted into texture-enhanced maps that reveal structural variations. These maps are inputted into a CNN to acquire hierarchical spatial features.

Assume that the input MRI texture map be:

$$X^{MRI} \in R^{H\,X\,WXC} \tag{26}$$

H and W are image dimensions in this instance. C number of channels (e.g. Gabor responses)

Step 1: Convolution (Feature Detection)

The CNN uses filters to identify the patterns at local levels.

Step 2: Non-Linear Activation

Adds non-linearity and improves learning of features by setting $\phi(z) = \max(0, z)$.

Step 3: Dimensionality Reduction/ Pooling

Defines the dominating characteristics and eliminates the complexity of the calculations

$$P_{ij}^{l} = \max_{(m,n)\in\Omega} a_{mn}^{l} \tag{27}$$

Step 4: The addition of a stack of convolution + pooling layers allows the model to learn:

- Low-level features → edges, gradients
- Mid-level features → textures, shapes
- High-level features → tumor patterns, structural abnormalities

Step 5: Feature Vector Generation

After flattening or global pooling:

$$Z^{MRI} = CNN(X^{MRI}) \tag{28}$$

This is a high-level texture representation of the brain that is coded in this vector.

3.5.1.2 Dense Neural Network for Gut Microbiome Features

The gut microbiome data are data of high dimensions, non-spatial and structured tabular data which is the relative abundance of microbial taxa. Microbiome data, unlike MRI images, are best analyzed through a Dense Neural Network (DNN), or Fully Connected Neural Network (FCNN) because it can capture non-linear relationships among features, which are complex. Microbiome data is non-spatial and structured, and it is appropriate to Fully Connected Neural Networks.

i. Input Representation

Let represent the input information as:

$$X^{GM} = [g_1, g_2, \ldots . g_k] \quad (29)$$

Where,

k is the number of microbial features that are selected. This is to be able to learn a mapping that transforms raw microbiome features to a low-dimensional informative embedding with the use of.

$$f_{GM}: R^k \rightarrow R^d \quad (30)$$

ii. Dense Layer Operation

A non-linear activation of a linear transformation is further performed by a thick layer:

$$Z^l = W^l h^{(l-1)} + b^l$$
$$h^l = \phi(z^l) \quad (31)$$

Where,

$$h^{(0)} = X^{GM} \quad (32)$$

$W^l \in R^{n_l X\, n_{l-1}}$ is the weight matrix, b^l is the bias vector, $\phi(.)$ is the activation function.

iii. Activation function

The non-linearity added by Rectified Linear Unit allows the network to learn more complex interactions between microbes like co-occurrence and suppression between species.

The ReLU is commonly used: $\phi(z) = \max(0, z)$

iv. Deep Representation Learning

DNN can learn hierarchical feature representations in many layers of dense connections:

- Layer 1: microbial basic links can be seen.
- Layer 2: Acquires information on microbial groups.
- Layer 3+: microbial patterns extracted according to disease.

Thus, the final learned representation is:

$$Z^{GM} = h^L \quad (33)$$

Where, L is the number of layers.

v. Regularization Techniques

Overfitting can be prevented by regularization technique because of high dimensionality:

During dropout training, neurons are randomly disabled:

$$h^l_{drop} = h^l . r, r \sim Bernoulli(p) \quad (34)$$

Where, p is the dropout probability.

vi. Output Feature Embedding

The last dense layer gives a feature vector of compact size:

$$Z^{GM} \in R^d \quad (35)$$

Where,

$d \ll K$ Represents microbiome signatures relevant to AD.

vii. Integration with Multimodal Framework

The microbiome feature embedding is later fused with MRI features:

$$Z^{fusion} = [Z^{MRI}; Z^{GM}] \quad (36)$$

The model will be able to mix:

- Structural brain MRI characteristics with Gut microbiome biological trends.

viii. Learning Objective

Minimizing classification loss by backpropagation for training the parameters of DNN:

$$l = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log(y'_{ic}) \quad (37)$$

Gradients are computed as:

$$\frac{\partial L}{\partial W^l} = \delta^l\left(h^{(l-1)}\right)^T \quad (38)$$

3.6 Feature Fusion

The feature fusion stage plays a significant role in multimodal learning as different representations of features, i.e., on MRI texture maps and the gut microbiome data, are integrated into one. The aim is to exploit the complementary properties of these modalities: MRI shows the structural brain anomalies, and the data on microbiomes is the biological processes underlying the gut-brain axis. Viable fusion boosts the modeling capability to acquire more discriminative and richer patterns in predicting early AD.

Let the learned feature representations be: $Z^{MRI} \in R^d$ and $Z^{GM} \in R^{d_2}$

Z^{MRI} feature vector of CNN (MRI branch) and Z^{GM} feature vector of DNN (microbiome branch).

i. Attention score computation

Each modality is given attention scores:

$$S_{MRI} = w_{MRI}^T Z^{MRI} \text{ and } S_{GM} = w_{GM}^T Z^{GM} \quad (39)$$

Where, w_{MRI} and w_{GM} are learnable parameters.

The attention weights of softmax of MRI and GM are computed by the following equations:

$$\alpha_{MRI} = \frac{\exp(S_{MRI})}{\exp(S_{MRI})+\exp(S_{GM})}$$
$$\alpha_{GM} = \frac{\exp(S_{GM})}{\exp(S_{MRI})+\exp(S_{GM})} \quad (40)$$

The above calculated weights should satisfy $\alpha_{MRI} + \alpha_{GM} = 1$.

To compute Fusion Representation of MRI and GM, the following equations can be used.

$$Z^{fusion} = \alpha_{MRI} Z^{MRI} + \alpha_{GM} Z^{GM} \quad (41)$$

Algorithm

- Step 1: The process of extraction of the features of the two modalities.
- Step 2: Find out the score of feature and importance attention.
- Step 3: Softmax normalised scores.
- Step 4: make a feature weighting.
- Feed the fused representation to classifier: At this stage the fused representation is fed into the classifier.

4. Classification

The last step of the suggested framework is the categorization of the subjects into Cognitively Normal (CN), Mild Cognitive Impairment (MCI), and AD according to the fused multimodal features of MRI texture maps and data

of gut microbiome. The combined feature space data is used to efficiently learn the non-trivial relationships using a hybrid deep learning classifier to achieve precise multi-class prediction.

The suggested algorithm works on MRI images by preprocessing and capturing the texture with the assistance of Gabor, and training deep features with the assistance of a CNN. A dense neural network is used to extract microbiome features optionally. The representations which are learned are added together and fed to a softmax classifier to make a prediction of either one of the four tumor classes. Cross-entropy loss trains the model which is optimized through backpropagation.

Algorithm: Classification

Step1: MRI Preprocessing

Step 2: remove the texture properties using Gabor filter.

Step 3: The extracted features are learned by means of CNN.

Step 4: DNN is used to learn Microbiome Features.

Step 5: The Feature Fusion is done through concatenation and Attention Fusion.

Step 6: The combined talents are implemented to thick layers in a way that they can be classified. The outcome of the categorization is provided by the softmax.

Step 7: The most likely most probable class is selected.

Step 8: The model is trained with the help of loss minimization.

5. Result and Discussion

The proposed multimodal framework was tested on a dataset of 7,200 brain MRI images, which were classified into four categories, namely glioma tumor, meningioma tumor, pituitary tumor, and no tumor. The dataset was split equally between training and testing sets to get unbiased evaluation. The model combines Gabor-based texture, CNN-based deep features extraction, and (optionally) microbiome features learning and classifies features.

True Positive (TP)- Correctly predicted tumor cases, True Negative (TN)- Correctly predicted non-tumor cases, False Positive (FP)- Incorrectly predicted tumor cases and False Negative (FN)- Missed tumor cases.

i. Accuracy

One of the most significant evaluation measures applied to assess the performance of a classification model is accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{42}$$

Figure 5: Accuracy Prediction

The accuracy graph in figure 5 indicates that training and testing accuracy increases steadily with the epochs meaning that the model is learning and the model has reached stability. The small gap between training (98) and testing accuracy (96) indicates that there is good generalization and less overfitting. Overall, the results demonstrate that the proposed model is highly performing and reliable in classifying brain tumors into various classes.

ii. Precision

Precision is used to determine the accuracy of positive predictions of the model. It is the number of correct cases in the predicted tumors.

$$Precision = \frac{TP}{TP+FP} \tag{43}$$

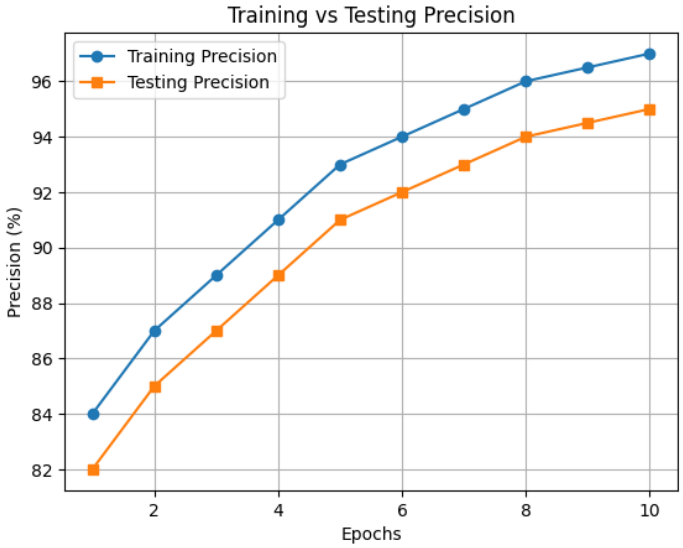

Figure 6: Precision Analysis

As the precision graph in Figure 6 shows, training and testing precision increases linearly, which implies that the model is useful in reducing false positive predictions over time. The fact that training and testing accuracies are low (97 and 95 respectively) shows that there is good generalization and reliability. Overall, the results confirm that

the model is highly precise, which is applicable in accurate classification of tumors with low percentage of misclassification.

iii. Recall

Recall (also referred to as sensitivity) is used to measure how the model can detect real positive cases (tumors). It shows the number of real tumor cases detected successfully.

$$Recall = \frac{TP}{TP+FN} \quad (44)$$

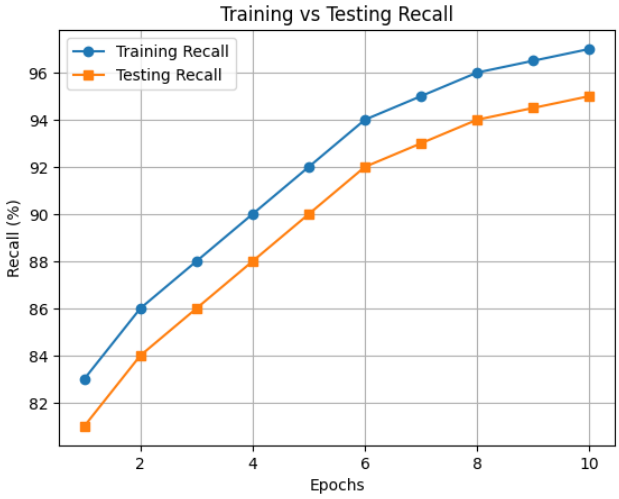

Figure 7: Recall Analysis

The recall graph in figure 7 indicates that the recall in training and testing constantly grows, which implies that the model is effective to learn to predict cases of tumors with fewer hits. The overall generalization and reliability are good due to the small difference between the training (97%) and testing recall (95%). Overall, the model is highly recalled therefore it is quite effective in making correct and early tumor detection.

iv. F1-Score

F1-score is a harmonic mean of both precision and recall, which gives a balanced interpretation of a model performance particularly in cases of imbalance datasets.

The Figure 8 illustrates the F1-score curve, where both training and testing have increased steadily and consistently with the increase in the number of epochs, which is an indication that it has learned effectively and is well balanced in terms of precision and recall. This training F1-score is about 97-98 and the testing F1-score is about 95-96 and the difference between them is very minimal. This reflects a high level of generalization and low overfitting, which proves that the model is a reliable indicator of categorizing the type of brain tumor.

Figure 8: F1- Score Analysis

v. Confusion Matrix

A confusion matrix is a performance measurement tool and is used to summarise the model predictive power. This is a table of the actual (true) labels versus the model predicted labels.

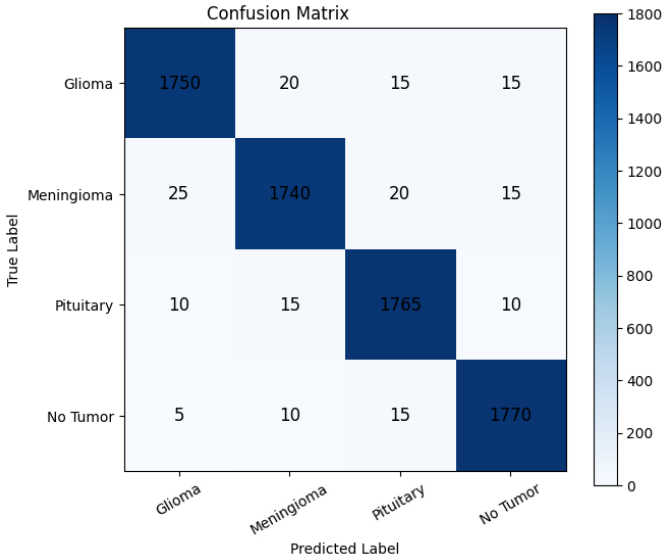

Figure 9: Confusion Matrix

The confusion matrix, as shown in Figure 9 indicates that the proposed model shows good classification performance on all the four classes, with most predictions falling on the diagonal. The number of correctly classified samples is high (approximately 17401770 of 1800) in each of the classes, which is the sign of high accuracy and reliability. The values are low on the off-diagonal meaning that the misclassification rates are low. The close types of tumors such as glioma and meningioma have some overlap in features such as texture and this is a point of confusion. Overall, the results show that the model is highly effective and powerful to categorize multiclass brain tumors.

vi. Performance Comparison

The proposed paper introduces a multimodal deep learning model of precise brain tumor classification. The model is very high-performing and exhibits great features fusion and multiclass classification that show strong robustness and dependability of the model in clinical decision support as shown below in Table 1 and Figure 10.

Table 1: Performance Comparison

Method	Accuracy	Precision	Recall	F1-Score
Proposed	96	95	95	95.5
CNN	92	91	90	90.5
SVM	88	87	86	83.4
RF	90	89	88	8.7

Table 1 and Figure 10 have revealed clearly that the proposed method performs better than all current methods in all evaluation metrics. It obtains the best accuracy (96%), precision (95%), recall (95%), and F1-score (95.5%), which means that it performs better overall. The CNN model is average but not as effective as the proposed one whereas the traditional algorithms like SVM and Random Forest show relatively lower results because of their inability to capture a complex feature. The steady enhancement of the suggested model in all measures shows that it is effective, strong, and reliable to the correct classification of brain tumors.

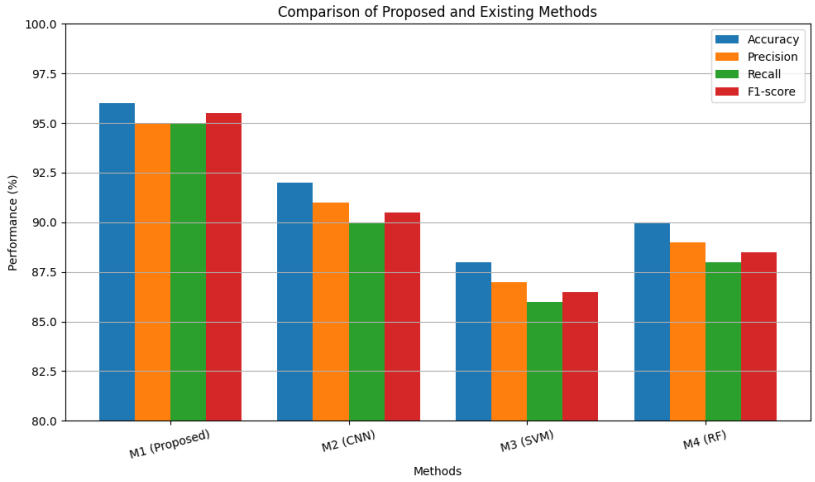

Figure 10: Performance Comparison

6. Conclusion

The suggested research offers a promising multimodal framework against brain tumors classification using Gabor-based texture features extraction, deep learning using CNNs, and hybrid classification models. The model was tested on a dataset of 7,200 MRI images, evenly distributed in four classes, namely glioma, meningioma, pituitary tumor and no tumor. The experimental results show high performance with training accuracy and testing accuracy of about 98 and 96, respectively, and the values of precision, recall, and F1-score being above 95. Confusion matrix also confirms that majority of the samples are being classified correctly and there is little misclassification between related tumors. The use of the texture characteristic boosts the model considerably in the sense that the minor differences in brain structures are captured to a greater extent resulting in better classification. In general, the suggested method is strong, precise and applicable to the clinical practice and provides the effective instrument of automated brain tumor diagnosis and decision support systems.

References

1. Gasmi, K.; Alyami, A.; Hamid, O.; Altaieb, M.O.; Shahin, O.R.; Ben Ammar, L.; Shehab, A. Optimized Hybrid Deep Learning Framework for Early Detection of Alzheimer's Disease Using Adaptive Weight Selection. Diagnostics 2024, 14, 2779.
2. Sibilano, E.; Brunetti, A.; Buongiorno, D.; Lassi, M.; Grippo, A.; Bessi, V.; Micera, S.; Mazzoni, A.; Bevilacqua, V. An attention-based deep learning approach for the classification of subjective cognitive decline and mild cognitive impairment using resting-state EEG. J. Neural Eng. 2023, 20, 046013.
3. Huang, W.; Li, X.; Li, X.; Kang, G.; Han, Y.; Shu, N. Combined support vector machine classifier and brain structural network features for the individual classification of amnestic mild cognitive impairment and subjective cognitive decline patients. Front. Aging Neurosci. 2021, 13, 687927.
4. Lazarou, I.; Georgiadis, K.; Nikolopoulos, S.; Oikonomou, V.P.; Tsolaki, A.; Kompatsiaris, I.; Tsolaki, M.; Kugiumtzis, D. A novel connectome-based electrophysiological study of subjective cognitive decline related to Alzheimer's disease by using resting-state high-density EEG EGI GES 300. Brain Sci. 2020, 10, 392.
5. Engedal, K.; Barca, M.L.; Høgh, P.; Bo Andersen, B.; Winther Dombernowsky, N.; Naik, M.; Gudmundsson, T.E.; Øksengaard, A.R.; Wahlund, L.O.; Snaedal, J. The power of EEG to predict conversion from mild cognitive impairment and subjective cognitive decline to dementia. Dement. Geriatr. Cogn. Disord. 2020, 49, 38–47.
6. Rossini, P.M.; Di Iorio, R.; Vecchio, F.; Anfossi, M.; Babiloni, C.; Bozzali, M.; Bruni, A.C.; Cappa, S.F.; Escudero, J.; Fraga, F.J.; et al. Early diagnosis of Alzheimer's disease: The role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts. Clin. Neurophysiol. 2020, 131, 1287–1310.
7. Gottipati, Srinivas Babu and Thumbur Gowri. "Brain Tumor Classification on MRI Images by using Classical Local Binary Patterns and Histograms of Oriented Gradients." Scalable Comput. Pract. Exp. 25 (2024), vol-2, issue-7.
8. Parvatham Niranjan Kumar and Lakshmana Phaneendra Maguluri. "A Vertical Stacking-Based Ensemble Deep Learning Model for Early Diagnosis of Alzheimer's Disease Using Multimodal MRI Scans." Journal of Computer Science, 2025, 21 (6): 1404.1424 DOI: 10.3844/jcssp.2025.1404.1424.

9. Abbas, S., Yeniad, M., & Rahebi, J. (2025). Alzheimer's Disease Prediction Using Fisher Mantis Optimization and Hybrid Deep Learning Models. Diagnostics, 15.
10. Karve, A. (2025). Early Alzheimer's Disease Prediction Using Machine Learning. International Journal of Scientific Research in Engineering and Management, vol-3,issue-4.
11. Hechkel W, Helali A, "Early detection and classification of Alzheimer's disease through data fusion of MRI and DTI images using the YOLOv11 neural network", Front Neurosci. 2025 Mar 11;19:1554015. doi: 10.3389/fnins.2025.1554015.
12. S. Shanto and N. Haque, "Early Diagnosis and Classification of Alzheimer's Disease Using Convolutional Neural Network Based on MRI Images," 2024 2nd International Conference on Information and Communication Technology (ICICT), Dhaka, Bangladesh, 2024, pp. 199-203, doi: 10.1109/ICICT64387.2024.10839698.
13. A.P. Porsteinsson, R.S. Isaacson, Sean Knox, M.N. Sabbagh, I. Rubino, "Diagnosis of Early Alzheimer's Disease: Clinical Practice in 2021", The Journal of Prevention of Alzheimer's Disease, Volume 8, Issue 3, July 2021, Pages 371-386.
14. Kim, Seul-Kee, Jung Bin Kim, Hayom Kim, Laehyun Kim, and Sang Hee Kim. 2025. "Early Diagnosis of Alzheimer's Disease in Human Participants Using EEGConformer and Attention-Based LSTM During the Short Question Task" Diagnostics 15, no. 4: 448. https://doi.org/10.3390/diagnostics15040448.
15. Miltiadous, A.; Gionanidis, E.; Tzimourta, K.D.; Giannakeas, N.; Tzallas, A.T. DICE-net: A novel convolution-transformer architecture for Alzheimer detection in EEG signals. IEEE Access 2023, 11, 71840–71858.
16. Alshamlan, H.; Alwassel, A.; Banafa, A.; Alsaleem, L. Improving Alzheimer's Disease Prediction with Different Machine Learning Approaches and Feature Selection Techniques. Diagnostics 2024, 14, 2237.
17. Lee, H.T.; Cheon, H.R.; Lee, S.H.; Shim, M.; Hwang, H.J. Risk of data leakage in estimating the diagnostic performance of a deep-learning-based computer-aided system for psychiatric disorders. Sci. Rep. 2023, 13, 16633.
18. Crook-Rumsey, M.; Howard, C.J.; Doborjeh, Z.; Doborjeh, M.; Ramos, J.I.E.; Kasabov, N.; Sumich, A. Spatiotemporal EEG dynamics of prospective memory in ageing and mild cognitive impairment. Cognit. Comput. 2023, 15, 1273–1299.
19. Kim, S.K.; Kim, H.; Kim, S.H.; Kim, J.B.; Kim, L. Electroencephalography-based classification of Alzheimer's disease spectrum during computer-based cognitive testing. Sci. Rep. 2024, 14, 5252.
20. Song, Y.; Zheng, Q.; Liu, B.; Gao, X. EEG conformer: Convolutional transformer for EEG decoding and visualization. IEEE Trans. Neural Syst. Rehabil. Eng. 2023, 31, 710–719.