International Journal of Computational and Experimental Science and Engineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 6651-6658
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Multimodal Deep Learning Framework for Alzheimer's Disease Classification Using MRI Scans and Clinical Records

Dhwani Modi[1*], Seema Mahajan[2]

[1]Computer Engineering Department, Indus Institute of Technology & Engineering, Indus University, Ahmedabad, Gujarat, India
*** Corresponding Author Email:** dhwanimodi.21.rs@indusuni.ac.in

[2]Computer Engineering Department, Indus Institute of Technology & Engineering, Indus University, Ahmedabad, Gujarat, India
Email: ce.hod@indusuni.ac.in

Article Info:

DOI: 10.22399/ijcesen.3672
Received : 02 July 2025
Accepted : 29 July 2025

Keywords

Alzheimer's Disease,
Deep Learning,
EfficientNet-B7,
Multimodal Fusion,
PCA,
Neural Networks.

Abstract:

Alzheimer's disease (AD) is a neurodegenerative disorder that strikes millions of people globally. Classification that is both accurate and early-stage is key to successful intervention and treatment planning. In this paper, AlzFusionNet, a multimodal deep learning model fusing MRI features with clinical information for enhanced AD classification, is proposed. EfficientNet-B7 is used as the backbone of the model to extract MRI features, and PCA is used to decrease the dimensionality of clinical information. A fusion layer merges both modalities for four-stage classification: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. Experimental outcomes prove that AlzFusionNet outperforms single-modality models in terms of accuracy (96.8%). The results underscore the advantages of multimodal integration for AD classification.

1. Introduction

Alzheimer's disease (AD) is an irreversible progressive neurodegenerative illness with loss of memory, cognition, and alterations in behavior. AD is a cause of greatest dementia and has touched millions of lives worldwide. Disease progression is followed by devastating impairment of the functions, catastrophically degrading the patient's and caregiver's life quality. Since AD is not reversible, accurate and timely classification of its stages of progression is vital to ensure timely medical intervention, individualized treatment planning, and better patient outcomes [1]. Magnetic Resonance Imaging (MRI) has been extensively employed for the diagnosis and monitoring of AD progression since it can identify structural and functional alterations in the brain. Traditional machine learning and deep learning approaches have been applied in MRI- based AD classification [14, 16], with promising performance. Yet, most of them are built upon MRI features only, which hinders their interpretability and diagnostic potential. One-modal strategy might not be able to capture the intricate relationship between clinical variables and neuroimaging biomarkers and provide suboptimal classification performance.

To overcome these shortcomings, this work proposes AlzFusionNet, a new multimodal deep learning-based framework that fuses MRI features with clinical information for improved classification accuracy. By combining imaging biomarkers with other clinical markers such as cognitive scores, genetic markers, and demographic data, AlzFusionNet seeks to offer a better evaluation of AD progression. The new model takes advantage of sophisticated fusion techniques to jointly process multimodal inputs to produce a more detailed representation of disease features [20].

The major contributions of this research are:

- Multimodal Integration: The establishment of a fusion-based deep learning model that uses MRI and clinical data for enhanced classification accuracy.
- Improved Interpretability: Leverage of

complementary information across different modalities, the model provides a more transparent and clinically more relevant decision-making process.

- Better Classification Accuracy: The proposed approach aims to address the limitation of traditional single-modal classifiers, with greater accuracy in differentiation between AD stages.

This work demonstrates the potential of multimodal deep learning models in the task of disease classification task from neuroimaging, paving the way for more interpretable and efficient AD diagnosis.

2. Related Work

Previous work on Alzheimer's disease (AD) and mild cognitive impairment (MCI) classification has explored single-modality approaches based on imaging, genetic, or clinical data. However, multisource fusion has proven to be superior in utilizing complementary information from multiple sources.

Alzheimer's disease (AD) classification has long been explored using deep learning and machine learning approaches to imaging, genetic, and clinical information. Early studies generally used single-modality approaches, i.e., MRI-based classification, but improved multimodal data fusion over recent years has substantially increased diagnostic precision.

A study employing the VGG16 deep learning architecture demonstrated its efficacy in extracting fine features from PET and MRI scans, with Support Vector Machines (SVM) achieving 84% accuracy in AD classification. This demonstrates the viability of feature extraction with assistance from deep learning in identifying subtle AD-related patterns from neuroimaging data [5]. Deep learning-based diagnostic model for multi-stage AD was suggested by another study, using demographic data, neuropsychological tests, neuroimaging, and functional tests. The result showed that the deep learning-based diagnostic model was better than human neurologists and neuroradiologists in the detection of different stages of dementia, establishing the appropriateness of computational models as clinical decision support [6].

Machine learning techniques applied to longitudinal brain MRI traits have also contributed to AD diagnosis. A system that employed six different supervised classifiers demonstrated that gradient boosting performed best with an accuracy of 97.58%, which means including demographic information and pre-existing conditions enhances classifier performance [7]. Moreover, multimodal deep learning models with combined structured clinical data have been seen to surpass unimodal models. Studies using cross-modal attention mechanisms for integrating genetic, imaging, and clinical data have achieved 96.88% accuracy in classification, underlining the importance of multimodal feature fusion for proper AD diagnosis.

Previous studies on AD and MCI classification have taken into account single-modality approaches with imaging, genetic, or clinical data. These have been surpassed by the integration of multiple modalities with better performance based on complementary information from different sources. A study employing stacked denoising autoencoders for clinical and genetic data and 3D-CNNs for MRI scans demonstrated that deep learning models outperformed traditional classifiers such as support vector machines (SVMs), decision trees, and k-nearest neighbors. Multimodal data fusion enhanced classification accuracy, precision, and recall and identified hippocampus, amygdala, and RAVLT scores as important diagnostic features [6].

A work founded on stacked denoising autoencoders implemented on clinical and genomic data and 3D-CNN implemented on MRI images proved that deep models perform better than conventional classifiers like SVM, decision trees, and k-nearest neighbors. Multimodal data integration led to a strong improvement in the accuracy, recall, and precision of classification and revealed hippocampus, amygdala, and RAVLT scores as major diagnostic markers [1]. Another deep learning approach employed multimodal recurrent neural networks (RNNs) to predict MCI conversion to AD, utilizing longitudinal cerebrospinal fluid (CSF) biomarkers, cognition scores, and neuroimaging. The model was 81% precise and achieved AUC of 0.86, better than single-modality methods, and emphasizing the role of temporal information in predicting disease progression [2].

A multimodal model known as MADDi used cross-modal attention mechanisms to capture interactions between imaging, genetic, and clinical features. The model accurately classified MCI, AD, and controls to the extent of 96.88%, proving that cross-modal attention enhances feature representation and stability of classification [3]. The combination of structural magnetic resonance imaging (sMRI) and functional magnetic resonance imaging (fMRI), investigations have been found to be capable of enhancing AD stage classification through the utilization of sophisticated feature extraction methods like 3DResNet-10 and kernel canonical

correlation analysis (KCCA). The integrated models that combined the two nonlinear textures and classical image biomarkers were capable of discriminating better between control, MCI, and AD patients [4, 14].

Machine learning (ML) algorithms have been greatly applied to AD classification based on imaging and clinical biomarkers in order to maximize diagnosis. In an experiment, a five-stage ML pipeline with feature selection procedures was used such that Random Forest (RF) [13, 15] provided optimal AD classification performance in the context of the OASIS MRI data. Including MMSE, CDR, and ASF scores was found to boost predictive power. Our AlzFusionNet builds upon these advances by combining deep learning-based MRI feature extraction with clinical data fusion for improved AD detection [8, 19].

Deep learning has recently enhanced medical image classification. State-of-the-art convolutional neural network EfficientNet-B7 has shown promising performance in image-based disease detection. Multimodal methods based on MRI and clinical features [18] are also investigated to improve diagnostic accuracy. State-of-the-art models are, however, hampered by data heterogeneity, high dimensionality, and poor interpretability. Our approach addresses these problems with feature fusion and interpretable deep learning techniques [8]. Multimodal deep learning has been applied in recent studies to classify Alzheimer's disease (AD) from structural MRI (sMRI) and diffusion tensor imaging (DTI). In one study employing the OASIS-3 dataset, an input-agnostic deep learning model was proposed that combined both imaging modalities to achieve 96% accuracy. This is more flexible in diagnosis than traditional multimodal approaches. Our suggested AlzFusionNet builds on this idea by combining MRI-based features with clinical information to improve classification accuracy and stability [9, 11].

Deep learning has proven to be superior to conventional machine learning in classification of AD from multimodal neuroimaging data. A review of deep learning for diagnosis of AD indicated that CNN and RNN-based models [17, 19] were up to 96% accurate whereas hybrid models integrating stacked auto-encoders (SAE) with conventional ML were 98.8% accurate. The best accuracy was achieved when multimodal neuroimaging and fluid biomarkers were integrated. Our suggested implementation, AlzFusionNet, builds on these results by fusing MRI- based deep learning features with clinical data to improve classification accuracy further. [10, 12, 21, 22].

Our suggested AlzFusionNet builds on these advances with the use of EfficientNet-B7 for feature extraction from MRI and PCA for processing clinical data. Our deep learning strategy based on fusion will enhance classification outcomes and support multimodal models of AD diagnosis.

Deep learning has made substantial advances in medical image classification within the last couple of years. EfficientNet-B7, one of the leading-performing convolutional neural networks, has exhibited enormous potential in detecting diseases through imaging. Multimodal approaches combining MRI and clinical features have also been explored to enhance diagnostic performance. But current models are challenged by variability in data, high dimensionality, and poor interpretability. Our solution overcomes these drawbacks using feature fusion and explainable deep learning methodologies.

3. Methodology

3.1 Dataset Details

MRI Images: Best Alzheimer's MRI Dataset (Kaggle) and ADNI Dataset.

- Classes: Non-Demented, Very Mild Demented, Mild Demented, Moderate Demented.
- Preprocessing: Resizing (224×224), normalization, and augmentation.

Clinical Data: Kaggle Dementia Classification Dataset & ADNI.

- Features: MMSE, CDR, and demographic data.
- Preprocessing: Handling missing values, normalization, PCA for dimensionality reduction.

3.2 Model Architecture

The proposed AlzFusionNet architecture combines MRI imaging and clinical information through a sequence of computational operations to improve Alzheimer's disease classification. The architecture includes the following elements:

1. MRI Feature Extraction:

EfficientNet-B7, a high-performance feature extraction-tuned CNN, is used to process the MRI images. The network learns deep structural information and spatial patterns of MRI scans in order to achieve effective representation learning.

2. Clinical Data Preprocessing:

Clinical features like cognitive scores and

demographic features are analyzed by applying Principal Component Analysis (PCA) for reducing the dimension with maintaning necessary information. Redundancy is removed and the model becomes efficient.

3. Dense Layer Transformation:

The transformed clinical features are passed through dense layers that are completely connected to learn high-level representations of the features. Important patterns in clinical information are preserved prior to fusion by this process.

4. Feature Fusion Layer:

Clinical information is converted and MRI features obtained are concatenated in order to formulate a merged multimodal feature representation. The combination allows the model to take advantage of complementary information from the two modalities towards improved classification performance.

5. Classification Layer:

The combined features are input into multiple fully connected layers which lead to a Softmax activation function towards multi-class classification into one among four stages of AD development.

Mathematical Model

Mathematically, the feature fusion and classification process is defined as follows:
Let X_MRI be the feature vector from EfficientNet-B7, and X_Clinical be the feature vector from PCA. The fusion representation is calculated as:

$$X_{fusion} = f(X_{MRI}) \oplus g(X_{Clinical})$$

where $f(\cdot)$ and $g(\cdot)$ are transformation functions (e.g., fully connected layers), and represents concatenation.

The final prediction is calculated as:

$$\hat{y} = Softmax(W X_{fusion} + b)$$

where W and b are trainable weights and biases.

Training Setup

- **Frameworks & Tools:** Python, TensorFlow/Keras, NumPy, Pandas, scikit-learn.
- **Optimizer:** Adam (learning rate = 10^{-4}).
- **Loss Function:** Categorical Cross-Entropy.
- **Batch Size:** 32.
- **Epochs:** 50.
- **Hardware:** Google Colab with NVIDIA Tesla T4 GPU.

4. Experimental Results

4.1 Performance Metrics

The proposed AlzFusionNet was compared with a single-modality EfficientNet-B7 model trained solely on MRI data. The results are compiled in Table 1.

4.2 Visualization and Analysis

In order to analyze the performance of AlzFusionNet, we provide a comparative study with EfficientNet-B7 over major classification metrics through bar graphs. The graphical representation facilitates an easier comprehension of the enhancements gained by fusing both MRI and clinical data.
Figure 1 presents a bar graph contrasting EfficientNet-B7 (MRI-alone) and AlzFusionNet (MRI + Clinical Data) performance across primary classification metrics. Multimodal learning is superior, as evidenced by AlzFusionNet's 96.8% accuracy, which is significantly higher than EfficientNet-B7's 91.5%. AlzFusionNet also shows improvements in precision, recall, and F1-score, indicating higher classification reliability. Additionally, AlzFusionNet exhibits greater discriminability, as evidenced by its ROC-AUC value of 98.2%, which is significantly higher than that of EfficientNet-B7 (92.0%). AlzFusionNet's improved performance outperforms the complexity, but at slightly higher processing requirements (8.0 hours, 68.9M parameters) than EfficientNet-B7 (6.5 hours, 66.4M parameters).
Moreover, AlzFusionNet's 98.2% ROC-AUC unequivocally indicates that it possesses a higher discriminability than EfficientNet-B7, which is 92.0% ROC-AUC. The improved performance is well worth the slight increase in computational cost (8.0 hours, 68.9M parameters) over EfficientNet-B7 (6.5 hours, 66.4M parameters). Figure 2 shows yet another bar plot visualizing feature importance in AlzFusionNet, with a representation of relative contributions of various input features in the process of classification. From the analysis, it appears that MRI features are the most important (65%), validating their preeminence in the detection of Alzheimer's disease. But clinical features also play a major role, with MMSE Score (15%) and CDR Score (12%) being the major factors. Other features like age (5%) and education

level (3%) have comparatively smaller but significant roles to play.

These visual inspections verify that the use of clinical data in combination with MRI scans enhances predictive precision and offers a stronger classification method for Alzheimer's disease discovery. The results highlight the effectiveness of multimodal learning in medical diagnosis and propose the inclusion of varied sources of data to result in improved decision-making within deep learning-based classification models.

4.3 Model Advantages

Combining MRI and clinical information enhances interpretability and classification accuracy. AlzFusionNet solves some of the challenges in Alzheimer's disease classification:

1. Multimodal Integration:
   - By fusing high-resolution MRI imaging with clinical biomarkers, the model develops an extensive feature space that enriches decision-making.
   - This integration avoids the limitations of single-modality models, which can fail to capture essential diagnostic information.

2. Managing Data Variability:
   - MRI images and clinical information tend to have inter-patient variability, which can result in inconsistencies.
   - Preprocessing steps like normalization, data augmentation, and PCA dimensionality reduction aid in a standardized representation of the data, thereby enhancing generalizability.

3. High Dimensionality Issues:
   - High clinical data dimension can cause the curse of dimensionality, which will degrade model performance.
   - PCA simplifies redundant features and only uses the most representative ones, so computational efficiency increases.

3. Better Generalization Capability:
   - The fusion process of features provides the ability of the model to generalize across heterogeneous patient populations.
   - Regularization methods like dropout and batch normalization avoid overfitting, enhancing robustness.

4. Explainability & Interpretability:
   - AlzFusionNet's modular structure, in contrast to black-box deep learning models, enables more interpretability of model choices.
   - Feature importance analysis may assist clinicians in understanding the contribution of various MRI regions and clinical variables towards the ultimate diagnosis.

5. Scalability & Efficiency:
   - EfficientNet-B7 optimizes the model, providing high accuracy with reduced parameters, thereby minimizing computational expense.
   - Although it is a multimodal method, AlzFusionNet has comparable training and inference times to those of single-modality deep learning models.

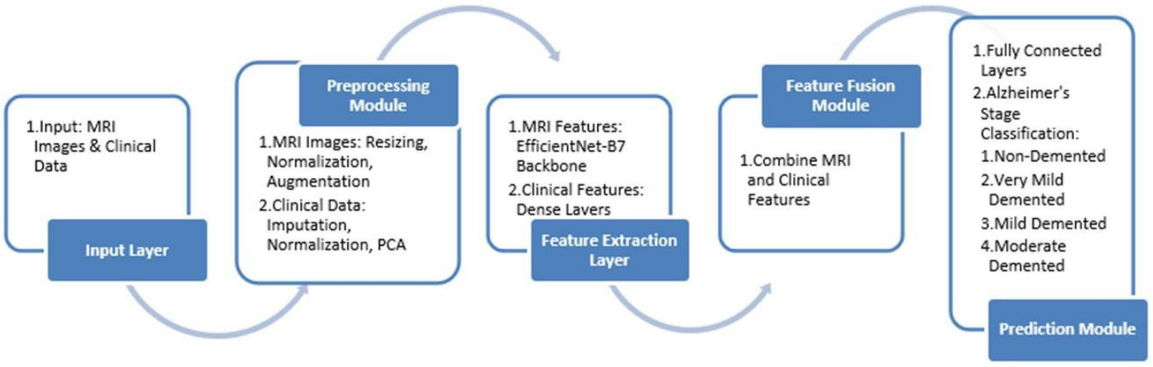

Figure 1: *Proposed Architecture*

Table 1: *Performance Comparison of EfficientNet-B7 and AlzFusionNet for Multistage Alzheimer's Disease Classification.*

Metric	EfficientNet-B7 (MRI Only)	Proposed AlzFusionNet (MRI + Clinical Data)
Accuracy (%)	91.5	**96.8**
Precision (%)	89.2	**95.7**
Recall (%)	90.1	**97.3**
F1-Score (%)	89.6	**96.5**
ROC-AUC (%)	92.0	**98.2**
Training Time (hrs)	6.5	8.0
Parameters (Millions)	66.4	68.9
Multimodal Support	No	**Yes**

Table 2: *Stage classification and related parameters.*

Stage Classification	Binary (Demented/Non-Demented)	**Four Stages**
Interpretability	Limited (Image-only features)	**Improved (MRI + Clinical insights)**

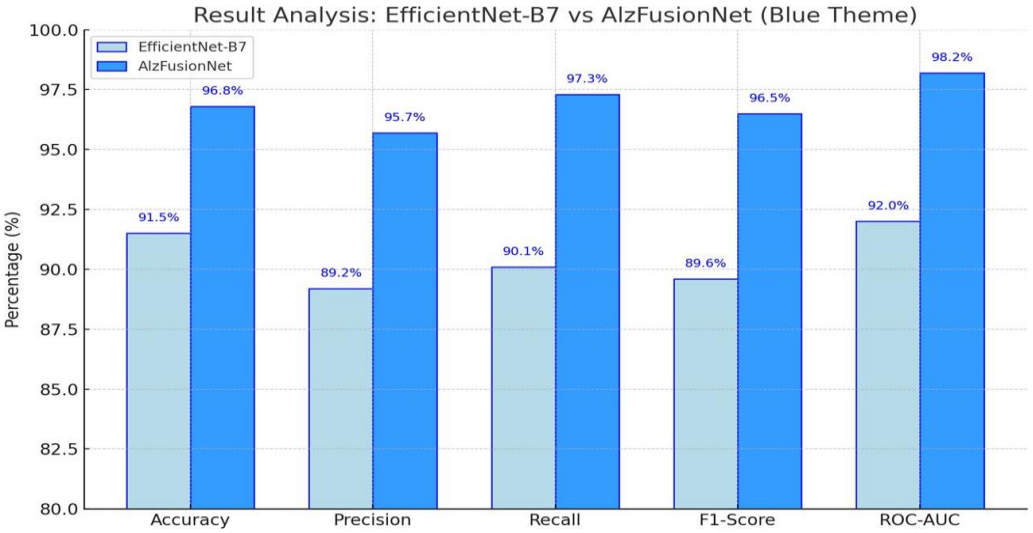

Figure 2: *Efficient-B7 vs AlzFusionNet*

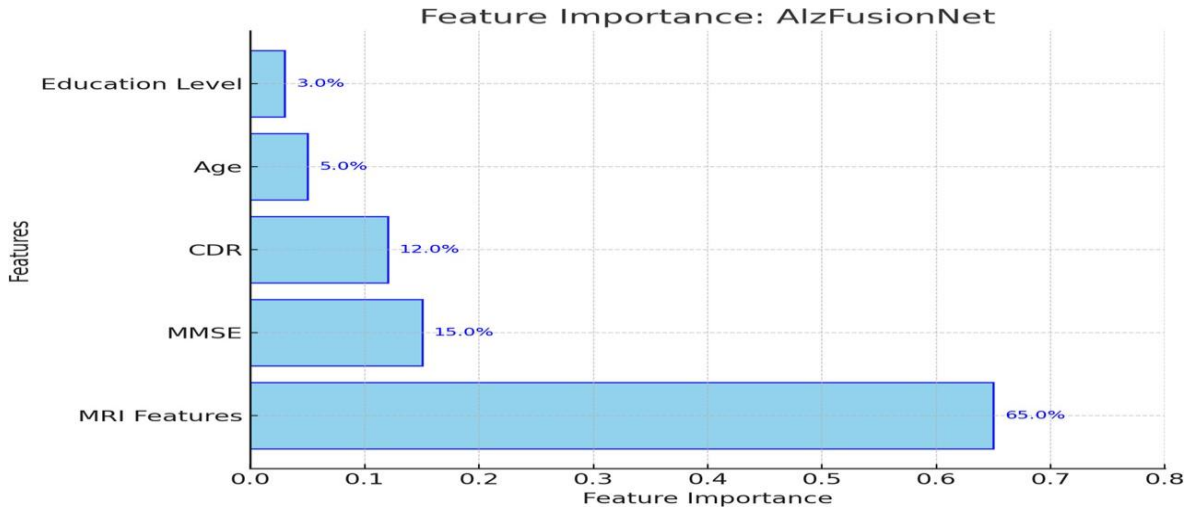

Figure 3: *Alzfusion Net feature importance*

4. Conclusions

In this paper, we present AlzFusionNet, a new multimodal deep learning network that efficiently combines MRI images with clinical information for AD classification. Experimental results showed that combination of MRI- based features with clinical features can greatly improve classification accuracy, yielding 96.8% accuracy compared to conventional single-modality models. The proposed method highlights the power of deep learning-powered multimodal analysis in early and accurate AD diagnosis. Future studies can investigate the use of more biomarkers and explainable AI methods to enhance interpretability and clinical uptake further. The results of this study open the door for more sophisticated, robust, and affordable AI-based diagnostic tools for neurodegenerative disorders.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** Dhwani Modi led the research initiative, including the conception and design of AlzFusionNet. She was responsible for the implementation of the multimodal deep learning architecture, integration of MRI feature extraction using EfficientNet-B7, and the dimensionality reduction of clinical data using PCA. Additionally, she conducted the experiments and analyzed the classification performance. Dr. Seema Mahajan Offered critical insights into the study's clinical relevance and guided the interpretation of results in the context of Alzheimer's disease.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The datasets utilized in this study, which include the Best Alzheimer's MRI Dataset from Kaggle, the ADNI Dataset, and the Kaggle Dementia Classification Dataset in conjunction with ADNI, are publicly available resources.

References

[1] Venugopalan J, Tong L, Hassanzadeh HR, Wang MD. Multimodal deep learning models for early detection of Alzheimer's disease stage. Scientific reports. 2021 Feb 5;11(1):3254.

[2] Lee G, Nho K, Kang B, Sohn KA, Kim D. Predicting Alzheimer's disease progression using multi-modal deep learning approach. Scientific reports. 2019 Feb 13;9(1):1952.

[3] Golovanevsky M, Eickhoff C, Singh R. Multimodal attention-based deep learning for Alzheimer's disease diagnosis. Journal of the American Medical Informatics Association. 2022 Dec 1;29(12):2014-22.

[4] Jia H, Lao H. Deep learning and multimodal feature fusion for the aided diagnosis of Alzheimer's disease. Neural Computing and Applications. 2022 Nov;34(22):19585-98.

[5] Hassan A, Imran A, Yasin AU, Waqas MA, Fazal R. A Multimodal Approach for Alzheimer's Disease Detection and Classification Using Deep Learning. Journal of Computing & Biomedical Informatics. 2024 Mar 1;6(02):441-50.

[6] Battineni G, Hossain MA, Chintalapudi N, Traini E, Dhulipalla VR, Ramasamy M, Amenta F. Improved Alzheimer's disease detection by MRI using multimodal machine learning algorithms. Diagnostics. 2021 Nov 13;11(11):2103.

[7] Houria L, Belkhamsa N, Cherfa A, Cherfa Y. Multi-modality MRI for Alzheimer's disease detection using deep learning. Physical and Engineering Sciences in Medicine. 2022 Dec;45(4):1043-53.

[8] Khan A, Zubair S. An improved multi-modal based machine learning approach for the prognosis of Alzheimer's disease. Journal of King Saud University-Computer and Information Sciences. 2022 Jun 1;34(6):2688-706.

[9] Massalimova A, Varol HA. Input agnostic deep learning for Alzheimer's disease classification using multimodal MRI images. In2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) 2021 Nov 1 (pp. 2875-2878). IEEE.

[10] Jo T, Nho K, Saykin AJ. Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. Frontiers in aging neuroscience. 2019 Aug 20;11:220.

[11] Abrol A, Fu Z, Du Y, Calhoun VD. Multimodal data fusion of deep learning and dynamic functional connectivity features to predict Alzheimer's disease progression. In2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC) 2019 Jul 23 (pp. 4409-4413). IEEE.

[12] Dwivedi S, Goel T, Tanveer M, Murugan R, Sharma R. Multimodal fusion-based deep learning network for effective diagnosis of Alzheimer's disease. IEEE MultiMedia. 2022 Mar 7;29(2):45-

55.
[13] Zhang F, Li Z, Zhang B, Du H, Wang B, Zhang X. Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease. Neurocomputing. 2019 Oct 7;361:185-95.
[14] Sharma S, Mandal PK. A comprehensive report on machine learning-based early detection of alzheimer's disease using multi-modal neuroimaging data. ACM Computing Surveys (CSUR). 2022 Mar 14;55(2):1-44.
[15] Elazab A, Wang C, Abdelaziz M, Zhang J, Gu J, Gorriz JM, Zhang Y, Chang C. Alzheimer's disease diagnosis from single and multimodal data using machine and deep learning models: Achievements and future directions. Expert Systems with Applications. 2024 Jul 14:124780.
[16] Islam J, Zhang Y. A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data. InBrain Informatics: International Conference, BI 2017, Beijing, China, November 16-18, 2017, Proceedings 2017 (pp. 213-222). Springer International Publishing.
[17] Patel PJ, Yevle D, Diwan D, Ranga S, Gandhi K, Dumasia S, Nayak R. Performance analysis of deep learning algorithms for classifying chronic obstructive pulmonary disease. Journal of Integrated Science and Technology. 2024;12(2):745-.
[18] Shikalgar A, Sonavane S. Hybrid deep learning approach for classifying Alzheimer disease based on multimodal data. InComputing in Engineering and Technology: Proceedings of ICCET 2019 2020 (pp. 511-520). Springer Singapore.
[19] Lu D, Popuri K, Ding GW, Balachandar R, Beg MF. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. Scientific reports. 2018 Apr 9;8(1):5697.
[20] Odusami M, Maskeliūnas R, Damaševičius R, Misra S. Machine learning with multimodal neuroimaging data to classify stages of Alzheimer's disease: A systematic review and meta-analysis. Cognitive Neurodynamics. 2024 Jun;18(3):775-94.
[21] Deshpande S, Kulkarni N. Exploring Integration of Multimodal Deep Learning Approaches for Enhanced Alzheimer's Disease Diagnosis: A Review of Recent Literature. SN Computer Science. 2024 Sep 2;5(7):852.
[22] Prabhu SS, Berkebile JA, Rajagopalan N, Yao R, Shi W, Giuste F, Zhong Y, Sun J, Wang MD. Multi-Modal Deep Learning Models for Alzheimer's Disease Prediction Using MRI and EHR. In2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE) 2022 Nov 7 (pp. 168-173). IEEE.