



OPEN MINDSETS: Multi-omics Integration with Neuroimaging for Dementia Subtyping and Effective Temporal Study

Salma Hassan^{1✉}, Dawlat Akaila¹, Maryam Arjemandi¹, Vijay Papineni² & Mohammad Yaqub¹

In the complex realm of cognitive disorders, Alzheimer's disease (AD) and vascular dementia (VaD) are the two most prevalent dementia types, presenting entangled symptoms yet requiring distinct treatment approaches. The crux of effective treatment in slowing neurodegeneration lies in early, accurate diagnosis, as this significantly assists doctors in determining the appropriate course of action. However, current diagnostic practices often delay VaD diagnosis, impeding timely intervention and adversely affecting patient prognosis. This paper presents an innovative multi-omics approach to accurately differentiate AD from VaD, achieving a diagnostic accuracy of 89.25%. The proposed method segments the longitudinal MRI scans and extracts advanced radiomics features. Subsequently, it synergistically integrates the radiomics features with an ensemble of clinical, cognitive, and genetic data to provide state-of-the-art diagnostic accuracy, setting a new benchmark in classification accuracy on a large public dataset. The paper's primary contribution is proposing a comprehensive methodology utilizing multi-omics data to provide a nuanced understanding of dementia subtypes. Additionally, the paper introduces an interpretable model to enhance clinical decision-making coupled with a novel model architecture for evaluating treatment efficacy. These advancements lay the groundwork for future work not only aimed at improving differential diagnosis but also mitigating and preventing the progression of dementia.

Keywords Neuroimaging, Dementia, Alzheimer's disease, Vascular dementia, MRI scans, Radiomics features, Multi-omics data, Brain segmentation

Dementia is a complex brain disorder, affecting 55 million individuals globally with ten million new cases annually, and represents a critical public health challenge¹. This umbrella term encompasses various cognitive disorders, including Alzheimer's disease (AD) and vascular dementia (VaD), which are the most prevalent forms. AD, responsible for 60-80% of dementia cases, is the sixth leading cause of death in the USA, with its mortality rates rising in contrast to those of heart disease and cancer^{2,3}. This trend underscores the urgent need for enhanced research, healthcare strategies, and societal support to address the escalating burden of AD. VaD, constituting 10-15% of dementia cases, shares symptoms with AD but requires distinct treatment due to its cerebrovascular pathology⁴. It results from cerebral ischemic or hemorrhagic events, leading to cognitive decline that varies with the affected brain regions. Early detection and management of vascular risk factors are critical for preventing VaD progression and improving patient outcomes.

The differential diagnosis between AD and VaD is vital given their unique etiologies and treatment approaches. AD primarily impairs memory and cognitive functions related to the hippocampus and entorhinal cortex, affecting language, navigation, and recognition skills⁵. In contrast, VaD affects executive functions early, resulting in difficulties with planning, organization, complex tasks, and mood disturbances^{6,7}. VaD patients have a more severe speech impairment and score up to 7.7% less on phonetic fluency tasks than AD patients⁷. The Hachinski Ischemic Score (HIS) is vital in distinguishing between degenerative and vascular dementia origins⁸.

Both AD and VaD have a profound impact on individuals' lives, extending far beyond the physiological symptoms. They affect patients' social, psychological, and economic well-being, but the nature and progression of cognitive decline differ between the two. Understanding the multifaceted nature of this impact emphasizes the critical importance of early diagnosis. Our research aims to bridge the gap in diagnostic capabilities, offering new insights and methodologies to enhance early detection and treatment strategies for AD and VaD.

¹Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, United Arab Emirates. ²Sheikh Shakhboub Medical City, Abu Dhabi, United Arab Emirates. ✉email: salma.hassan@mbzuai.ac.ae

Contributions: This paper introduces a method for differentiating AD and VaD using multi-omics data. The main contributions are:

- Propose a unified and robust multi-omics method to distinguish between AD and VaD from MRI radiomics, genetic data, and clinical assessments to reveal their pathological differences while achieving state-of-the-art classification performance on a large public dataset.
- Introduce an explainable model to facilitate its adoption in clinical practice and provide a personalized interpretable diagnosis.
- Develop a model architecture to monitor longitudinal changes in MCI diagnosis confidence and probability to assess treatment efficacy.

These contributions underscore our paper's potential to transform the diagnostic landscape for AD and VaD, offering a clinically applicable solution to bridge current gaps in the diagnosis and management of dementia.

Background

Diagnostic challenges in AD and VaD

Symptomatic overlap and clinical complications

The differential diagnosis of AD and VaD remains a formidable challenge in the neurocognitive domain, primarily due to the significant symptomatic overlap between these disorders. Both AD and VaD share common symptoms, such as memory loss, executive dysfunction, and difficulties with daily activities, which complicate clinical assessments⁹. This overlap complicates clinical evaluations and underscores the urgent need for precision in diagnosis to enable effective treatments and improve patient prognoses.

Limitations of traditional diagnostic tools

Traditional diagnostic tools, including clinical interviews, cognitive testing, and basic imaging techniques, often fall short in accurately differentiating between AD and VaD¹⁰. Cognitive assessments may not sufficiently capture the subtle differences in disease pathology, while standard imaging techniques like CT and conventional MRI may not provide the resolution needed to distinguish between the vascular and neurodegenerative changes characteristic of these conditions¹⁰. This inadequacy in current diagnostic methods highlights the necessity for more advanced, nuanced approaches.

Importance of accurate differentiation

Accurate differentiation between AD and VaD is crucial for several reasons. First, it ensures that patients receive the most appropriate treatments, which can vary significantly between the two conditions¹¹. AD treatments often focus on managing symptoms and slowing disease progression, while VaD management emphasizes preventing further vascular damage through aggressive control of cardiovascular risk factors. Second, precise diagnosis aids in better patient management and prognosis, allowing for tailored therapeutic strategies and more informed clinical decisions¹¹. Finally, it enhances our understanding of the pathophysiology of these dementias, contributing to the development of more effective interventions.

Differentiating factors between AD and VaD

AD typically has a gradual and progressive onset, characterized by a slow cognitive and functional decline with early loss of awareness¹². In contrast, VaD often presents with sudden onset due to cerebrovascular events, followed by a stepwise progression and focal neurological signs. Regarding attention and executive function, AD and VaD are equally impaired in sustained attention. However, selective attention is more impaired in VaD than AD, which can be attributed to cerebrovascular damage affecting brain regions responsible for executive functions¹². Regarding language assessment, semantic fluency is more impaired in AD, with patients struggling with naming and language comprehension tasks. In contrast, individuals with VaD perform worse on phonemic fluency measures, reflecting the impact of cerebrovascular damage on language processing areas¹³. Finally, regarding memory impact, AD shows greater overall memory impairment than VaD, particularly in episodic memory. Participants with VaD often perform better on episodic memory tests, while individuals with AD show significant deficits in recall memory tasks, which are crucial for diagnosis¹².

Vascular and neurodegenerative changes

VaD is primarily caused by tissue injury from ischemia, affecting the brain's white matter and subcortical structures. Commonly impacted areas include the basal ganglia, thalamus, and regions responsible for executive function and processing speed¹⁴. Ischemic lesions, small vessel disease leading to white matter hyperintensities, lacunar infarcts, and microbleeds are also present. On the other hand, AD is characterized by generalized brain atrophy, particularly in the hippocampus and temporal lobes, with early damage often seen in the hippocampus and entorhinal cortex¹⁵. As the disease progresses, the cerebral cortex, involved in thought, perception, and language, is affected, leading to ventricular enlargement and cortical thinning¹⁴. These distinct differences between the two types impact the treatment approaches. While AD treatment focuses on managing symptoms and slowing progression, VaD treatment emphasizes aggressive management of cardiovascular risk factors to prevent further vascular damage. The life expectancy of VaD is shorter, and its diagnosis often comes too late for effective recuperative treatment.

Advancements in machine learning and neuroimaging

Research utilizing the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset has significantly advanced our understanding of AD and VaD. Studies such as¹⁶ have demonstrated the potential of machine learning algorithms

combined with longitudinal structural MRI biomarkers to enhance diagnostic precision. These advancements offer valuable tools for clinicians, aiding in the early diagnosis of these distinct dementia types by identifying subtle brain changes over time. MRI provides detailed images of brain structure, allowing for the identification of dementia-specific brain changes. Studies have shown that MRI can detect subtle neurodegenerative markers such as white matter hyperintensities and lacunes, which are critical for distinguishing between AD and VaD in the early stages of disease progression^{17–19}.

Machine learning techniques

Machine learning techniques, including Support Vector Machines (SVM), K-nearest neighbors (KNN), and Random Forest (RF) classifiers, have improved diagnostic precision by analyzing complex patterns in neuroimaging and biomarker data. These techniques can identify distinctive brain volume differences and other biomarkers, facilitating more accurate early-stage dementia diagnosis^{16,20–22}.

Multimodal prediction in dementia

Recent advancements in AI-driven multimodal prediction for dementia have demonstrated promising results in enhancing early detection and risk assessment. The authors in²³ emphasized the importance of integrating neuropsychological testing, MRI imaging, and clinical risk factors in AI models, highlighting the role of dynamic frailty assessment for accurate late-life dementia prediction. Similarly, Vaghari et al.²⁴ validated the clinical utility of an AI-guided prognostic model that combines multimodal biological and cognitive data to predict cognitive decline. Furthermore, a different study highlighted the effectiveness of machine learning frameworks in predicting dementia progression in mild cognitive impairment (MCI) patients using cognitive tests, MRI, and PET scans²⁵. These studies collectively reinforce the potential of AI-driven multimodal approaches in dementia prediction, enhancing early intervention strategies and improving diagnostic accuracy.

Data fusion strategy

Recent studies on fusion strategies for multimodal prediction in dementia emphasize the importance of integrating diverse data sources to enhance predictive accuracy. The authors in²⁶ employed a data-level fusion approach, combining clinical data, MRI segmentation, and psychological assessments to improve Alzheimer's disease classification accuracy, leveraging Random Forest (RF) classifiers and SHapley Additive exPlanations (SHAP) for interpretability. Similarly, Ding et al. (2023) demonstrated feature-level fusion using neuropsychological tests, MRI biomarkers, and clinical risk factors, optimized through CatBoost with Optuna hyperparameter tuning, achieving an AUC of 0.90 for 10-year dementia risk prediction²⁷. Another novel approach utilized a deep learning-based three-stage fusion framework, where modality-specific latent representations were first extracted, then jointly learned at the feature level, and finally combined at the decision level for multi-status dementia diagnosis²⁸. Additionally, a new study explored hybrid fusion strategies, demonstrating that genetic data enhanced early dementia prediction for cognitively normal individuals, while MRI data provided better characterization for mild cognitive impairment patients²⁹. These studies highlight the significance of advanced fusion strategies-ranging from early feature-level fusion to late decision-level aggregation-to improve the robustness and interpretability of AI-driven dementia prediction models.

Multi-omics approach for improved diagnosis

In light of these developments, a multi-omics diagnostic framework is indispensable for addressing the complex interplay of symptoms in AD and VaD and deepening the understanding of their distinct pathophysiological traits. Therefore, we propose a novel approach, MINDSETS (Multi-omics Integration with Neuroimaging for Dementia Subtyping and Effective Temporal Study), to address this challenge. This approach leverages the integration of genetic, proteomic, and neuroimaging data to create a comprehensive diagnostic model that can more accurately distinguish between AD and VaD. By combining these diverse data sources, MINDSETS aims to improve diagnostic accuracy, enable early detection, and tailor treatment strategies to individual patients' needs.

Proposed MINDSETS approach

MRI pre-processing

Prior to feature extraction, all MRI scans underwent a comprehensive pre-processing pipeline using standard neuroimaging packages. First, N4ITK bias field correction was applied to address intensity inhomogeneities using ANTs (Advanced Normalization Tools). Brain extraction was performed using FSL's BET (Brain Extraction Tool) to remove non-brain tissues, followed by intensity normalization to ensure consistent intensity ranges across all scans. Spatial normalization was achieved through linear registration using FSL's FLIRT (FMRIB's Linear Image Registration Tool) and non-linear registration using ANTs' SyN algorithm to align all images to the MNI152 standard space. This registration process is crucial for ensuring spatial correspondence across different timepoints and subjects. The normalized images were then quality-checked to ensure successful pre-processing before proceeding with segmentation. To accommodate our dataset's characteristics, particularly the elderly and diseased population, we opted for a UNet-based segmentation model pre-trained on similar cohorts, which demonstrated robust performance across varying image resolutions and contrast levels.

MRI scan segmentation

The next step is to segment the brain to extract the radiomics features using the mask as illustrated in Fig. 1A. There are no ground truth segmentation masks available in the dataset. Thus, a pre-trained 3D UNet model was leveraged through SynthSeg, a deep learning tool adept at automatic brain MRI scan segmentation, notable for its robustness to contrast and resolution variations, trained across a diverse set of scans, including those from both elderly diseased and healthy populations³⁰. This allows the model to provide accurate estimation, even with

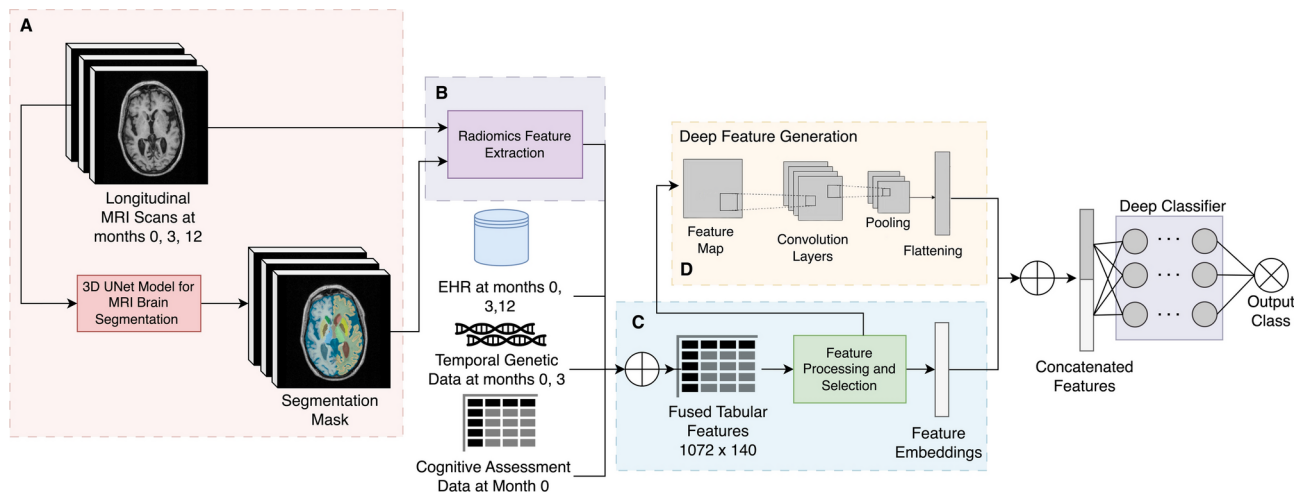


Fig. 1. The architecture segments the MRI scans to extract radiomics features and fuses them with other multi-omics features. Feature selection is applied, and then features are passed to the Deep Feature Generation module, where discriminative features are generated and concatenated with the raw features.

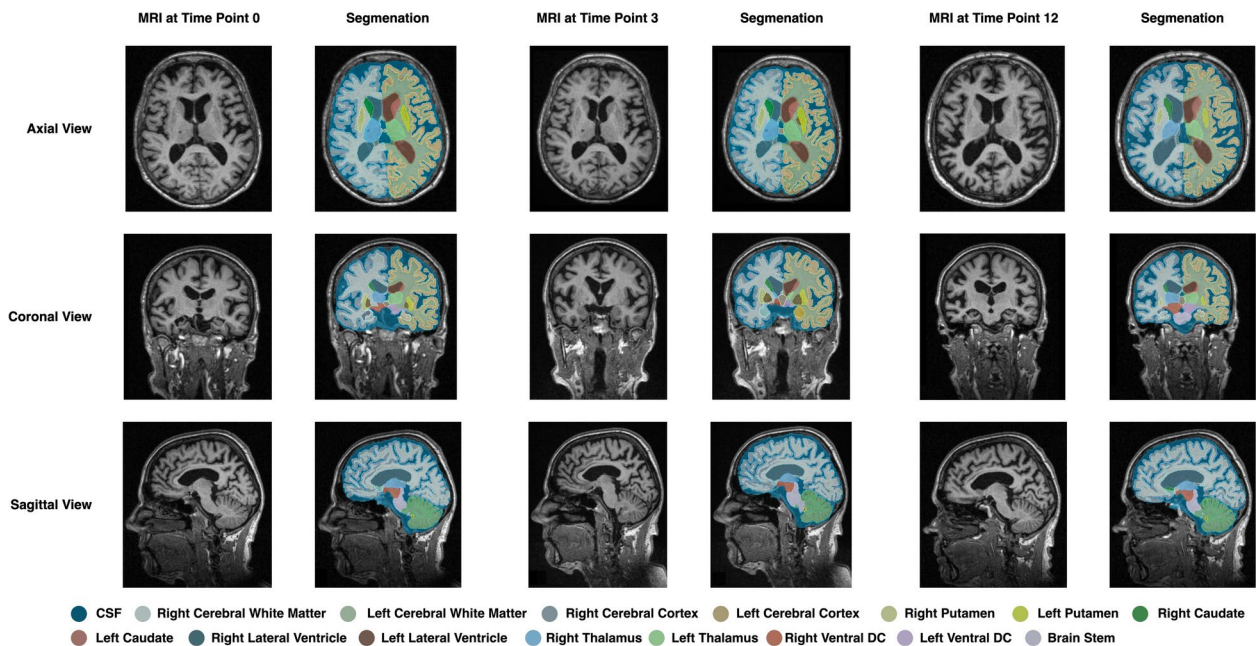


Fig. 2. Multi-planar MRI Scans with Annotated Segmentation: The image series showcases axial, coronal, and sagittal views of the temporal brain MRIs at time points 0, 3, 12, both in raw form (left column) and with detailed segmentation overlays (right column). The segmentation highlights various brain structures, including the thalamus and cortical regions, using distinct color codes to differentiate between anatomical areas.

brain atrophy present, which standard models such as FSL and FreeSurfer struggle to do³¹. This model segmented all the temporal MRI scans into 32 different structures, as illustrated in Fig. 2. Three criteria were utilized to validate the accuracy of these segmentation masks. Firstly, SynthSeg's automated quality control mechanism confirmed a high similarity score for all structures when compared to a standard reference. Secondly, comparing the volumes of the segmented structures against known true values revealed minimal discrepancies. Lastly, a manual assessment by an experienced radiologist on a random subset of scans was performed to assess the segmentation performance.

Radiomics features extraction

The next step was to utilize the segmentation mask for MRI feature extraction, as seen in Fig. 1B. To facilitate this, PyRadiomics, a Python package for radiomics analysis, was used³². This process involves extracting 137 distinct features from each of the 32 brain structures identified during segmentation, covering features related

to shape, first-order statistics, and gray-level attributes. Recent studies underscore the enhanced diagnostic accuracy achieved by using radiomics features, stressing its potential in quantitatively analyzing tissue patterns at the voxel level^{33,34}. This approach is particularly pertinent for distinguishing between AD and VaD, as these conditions manifest through specific changes in brain tissues.

Features fusion and selection

Following the extraction of MRI radiomics features, since the output of PyRadiomics is in tabular form, it was concatenated with the tabular clinical, genotype data and the mini-mental state examination (MMSE) cognitive assessment score, as shown in Fig. 1C. Since AD affects memory more prominently than VaD, the MMSE score was split into memory and processing-related sub-scores. After fusion, various feature selection approaches were compared. These included statistical methods such as chi-squared, principal component analysis, and t-distributed stochastic neighbor embedding. However, FeatureWiz provided the best result by reducing multicollinearity.

Deep feature generation

We propose the Deep Feature Generation (DFG) module to enhance the discriminative power of the architecture and surpass the current diagnostic accuracy, shown in Fig. 1D. The DFG module utilizes Convolutional Neural Networks (CNNs) to dynamically extract intricate patterns, thus enriching the representation of the data to combine local patterns into new features with deeper insights into the data. The new features are then concatenated with the raw features and passed to a deep classifier to output the class. This process empowers our architecture to capture nuanced relationships and subtle variations within the dataset, consequently improving its ability to discern relevant information for subsequent analysis. Through deep feature generation, we aim to augment the efficacy and robustness of our model, thereby advancing the accuracy and reliability of detention subtype classification.

One of the main aims of this paper is to provide explainability in terms of which features are most significant in deciding for diagnosis. Explainability is especially important in the medical field, where identifying critical features can provide vital insights into the diagnosis, treatment, and prevention of various medical conditions. To address this, several steps were implemented throughout the methodology to steer away from treating the model as a black box. For instance, feature importance scores were calculated during feature selection using techniques such as Random Forest and SHAP (Shapley Additive Explanations). This approach helped identify the most influential features and understand the direction and magnitude of their impact on the diagnostic predictions. Furthermore, visualizations were created to present these feature importance scores, allowing medical professionals to interpret them easily.

Explainability

Explainability is especially important in the medical field, where identifying critical features can provide vital insights into the diagnosis, treatment, and prevention of various medical conditions. To address this, several steps were implemented throughout the methodology to avoid treating the model as a black box. For instance, feature importance scores were calculated during feature selection using techniques such as Random Forest and SHAP (Shapley Additive Explanations). This approach helped identify the most influential features and understand the direction and magnitude of their impact on the diagnostic predictions. Furthermore, visualizations were created to present these feature importance scores, allowing medical professionals to interpret them easily.

Experimental setup

Dataset

Building a large-scale multi-omics dataset targeting dementia subtypes is expensive and time-consuming. There is only one public dataset to our knowledge that offers a rich foundation for dementia research, which is ANMerge³⁵ composed of 1,702 participants. Unlike the ADNI dataset, it encompasses EHR, longitudinal MRI scans, and comprehensive multi-omics data such as genomic profiles; example data is shown in Fig. 3. The genomic data in ANMerge is distributed in a preprocessed format as part of the standardized dataset release, and we utilized these features as provided to ensure both the reproducibility of our work and fair comparisons with other studies analyzing the same dataset. ANMerge’s emphasis on temporal and patient-level data, such as CDR and Camdex, enhances the model’s ability to analyze dementia progression and treatment efficacy, which is a unique advantage for personalized analysis compared to ADNI. There are four classes: AD, VaD, Mild Cognitive Impairment (MCI), and Control (CTL). This study leverages the dataset’s longitudinal MRI scans to critically discern degenerative versus vascular brain changes. Furthermore, it explores the diagnostic enhancement

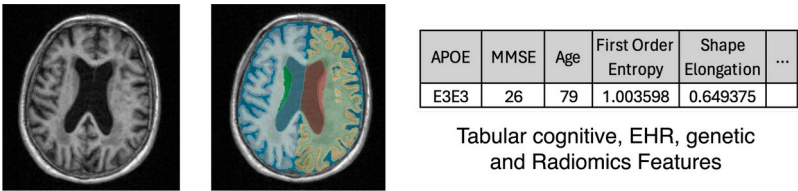


Fig. 3. An example patient data with MRI, the generated segmentation mask, and a sample set of the tabular data, left to right.

potential of integrating additional modalities - clinical data, cognitive assessments, and genetic information. Through experimentation, we assess the incremental value of each modality in improving diagnostic accuracy. Demographic and sample clinical features are summarized in Table 1.

Implementation details

For training, a 5 Group K-fold cross-validation strategy was implemented where data was split by patients to avoid leakage. Categorical features were encoded using label encoding, and continuous features were scaled using Standard Scaler. Missing information was imputed with the mean or mode in the case of continuous and categorical, respectively, based on the class group. The hyperparameters used were a maximum of 500 epochs but with early stopping implemented with a patience level of 5. For the DFG module, 14 filters were used with kernel size 7x7. The max pooling layer width was set to 2. The Optuna library was used for hyperparameters tuning³⁶.

Ablation studies

3D CNN model using MRI only

The first approach considered for this paper was an alternative to implementing the segmentation and radiomics method. Therefore, a direct image-based 3D CNN model was tested to investigate whether using raw 3D MRI alone could yield results comparable to our proposed methodology. Moreover, several machine learning and deep learning algorithms, such as random forest, support vector machine, multi-layer perception, and deep factorization machines, were implemented to test their performance against the proposed approach. Also, the model's performance was evaluated with and without the DFG module to isolate its contribution to the overall improvement.

Feature selection

Various feature selection approaches were implemented to compare. These included statistical methods such as chi-squared, variance threshold, principal component analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE), forward selection, backward elimination, and exhaustive searching. In addition, several libraries, such as mlxtend, were used. However, as mentioned, FeatureWiz provided the best result by working to reduce multicollinearity between the features while maintaining a high correlation with the target.

Model selection

Several ML and DL models were investigated to determine which would work best. Lazypredict library was used to quickly assess which ML algorithm would work best before hyperparameter tuning. SVM and random forest were the top two algorithms, respectively, with SVM outperforming RF in all experiments. A simple 1D CNN and tabular ResNet were implemented and tested for the DL approach. ResNet model provided a marginal boost in accuracy in some tests but was inconsistent through all the tests. TabNet library was also used to test its classifier, but it was not the highest performing. Finally, using the DeepTables library, all the model architectures available were tested, including DCN, DeepFM, FGCNN, and PNN. Finally, we implemented the proposed DFG module, which best performed among all other models tested.

MRI temporal analysis

Since the dataset contains longitudinal MRI scans, a crucial question is determining the specific time point at which the model excels in accurately diagnosing the dementia subtype. Thus, the experiment was repeated several times, including data at month 0 (baseline), month 3, and month 12 only to assess the MRI scan's importance and model adaptability for early differential diagnosis.

Multi-omics data integration

Finally, since multi-omics data are available, it is essential to investigate if their integration would improve the diagnosis accuracy by comparing results using multi-omics versus MRI only to account for real-life contexts

Features	AD (612)	MCI (425)	CN (425)	VaD (240)	Total (1702)
Sex					
Male # (%)	310 (50.65)	134 (31.52)	124 (29.18)	67 (27.92)	635 (37.31)
Female # (%)	302 (49.35)	291 (68.48)	301 (70.82)	173 (72.08)	1067 (62.69)
Age (years)					
Mean (std)	75.45 (6.60)	73.96 (5.74)	72.54 (6.64)	71.86 (6.32)	73.98 (6.4)
Range	[58,88]	[56,86]	[52,87]	[52,88]	[52,88]
Education (years)					
Mean (std)	7.99 (3.98)	8.97 (4.29)	11.01 (4.88)	10.85 (5.61)	9.32 (4.56)
Range	[2,22]	[0,20]	[2,25]	[2,25]	[0,25]
MMSE					
Mean (std)	20.80 (4.67)	27.09 (1.72)	29.07 (1.2)	22.60 (5.28)	25.71 (4.54)
Range	[12,30]	[24,30]	[25,30]	[14,30]	[12,30]

Table 1. Demographic and clinical features by diagnosis group.

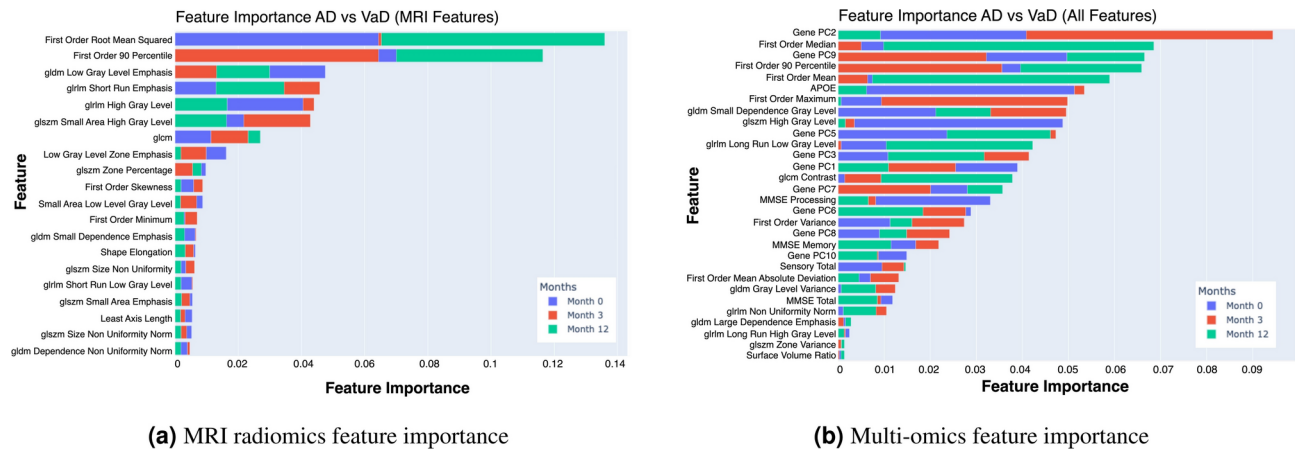


Fig. 4. Figure shows the feature importance of MRI radiomics features only versus multi-omics features. Importance is broken down by scan timepoint.

Data	Metric	AD Vs CTL		AD Vs MCI		MCI Vs CTL		AD Vs VaD		All 4 Classes	
		All	MRI at time 0	All	MRI at time 0	All	MRI at time 0	All	MRI at time 0	All	MRI at time 0
MRI	Accuracy	97.89	98.64	82.86	87.60	83.01	83.32	82.52	87.60	52.39	62.43
	F1-Score	97.86	98.59	82.44	87.41	81.78	81.84	72.36	79.03	48.32	54.58
	Recall	98.26	98.89	82.14	88.83	80.93	83.28	76.29	82.64	58.40	61.38
	Precision	97.65	98.33	83.20	87.25	81.95	81.04	70.35	76.61	58.55	67.64
	AUC	98.16	98.89	85.30	88.83	83.93	83.24	76.29	82.64	61.78	70.04
Multi-omics	Accuracy	99.35	99.92	89.52	90.69	83.89	86.45	88.60	89.25	70.12	72.23
	F1-Score	99.31	99.92	89.27	90.68	81.88	84.23	63.14	81.13	70.19	68.29
	Recall	99.33	99.91	90.24	90.69	81.24	82.63	59.02	83.61	75.25	66.77
	Precision	99.29	99.92	88.89	90.85	82.75	87.24	70.92	79.21	67.96	75.62
	AUC	99.33	99.91	90.69	90.24	81.24	82.63	77.07	83.61	71.26	77.32

Table 2. Summary of result using all longitudinal MRI scans, and scan 0 alone versus using it in with other multi-omics data with the DFG module. Results are shown for the binary groups tested and the multiclass experiment. Significant values are in bold, highlighting MINDSETS performance in early diagnosis using baseline scans only, in differential diagnosis and multi-class detection.

where genetic testing is expensive and not widely available, as can be seen from Fig. 4, which investigates the importance of the features.

Results and discussion
MINDSETS approach results

Our multi-omics approach integrates radiomics MRI features that were extracted from the segmentation mask of 32 brain structures. The evaluation showcased that 100% of the 30 randomly reviewed scans were consistent with the radiologist’s assessment. Quantitative results with the proposed DFG module are shown in Table 2 and without the DFG in Table 3 to compare its effect.

Qualitative results

A significant trend to notice is that based on the result, we can conclude that AD vs. CTL is the easiest group to differentiate, followed equally by AD vs. MCI and MCI vs. CTL, and finally, AD vs. VaD. This trend follows the same pattern as previously published work. In addition, this trend does adhere to the expected medical evaluation wherein it is easier to diagnose between a healthy patient and a patient with full onset AD with higher confidence, followed by the in-between cases where a healthy patient can be showing some cognitive decline that is harder to differentiate from typical aging symptoms or quantifying the measure of the cognitive decline to distinguish between MCI and AD. Since it is established that AD and VaD share a lot of similarities in terms of symptoms, it is expected to be the most challenging class to differentiate.

The qualitative aspects of utilizing baseline versus longitudinal MRI data reveal critical insights into the progression and differentiation of dementia subtypes. The baseline MRI provides a snapshot of the brain at a single point in time, which, while helpful, may not capture the dynamic nature of neurodegenerative diseases. In contrast, longitudinal MRI data, encompassing scans from months 0, 3, and 12, offer a temporal dimension that is invaluable for observing the trajectory of the disease. From a qualitative standpoint, the longitudinal approach

Data	Metric	AD Vs CTL	AD Vs MCI	MCI Vs CTL	AD Vs VaD	All 4 Classes
MRI	Accuracy	60.66	67.58	69.94	61.88	44.98
	F1-Score	60.48	67.31	67.27	58.90	40.64
	Recall	60.72	67.82	67.13	65.64	41.15
	Precision	60.95	68.60	67.46	76.70	49.04
	AUC	71.45	68.60	71.45	78.29	46.25
Multi-omics	Accuracy	81.27	80.79	76.45	69.72	62.91
	F1-Score	80.92	80.05	73.28	64.35	55.34
	Recall	80.98	80.22	72.66	76.94	56.69
	Precision	80.86	79.91	74.27	65.85	56.29
	AUC	80.98	80.22	73.66	78.94	58.54

Table 3. MINDSETS Approach Without Deep Feature Generation (DFG) Module—Table summarizing the evaluation metric using MRI scans alone versus with multi-omics data using the MINDSETS approach but omitting the DFG module.

allows for the detection of subtle changes and trends in brain morphology that could be indicative of early or progressive stages of dementia. This is particularly important in conditions like MCI, which may show minimal changes at month 0 but more pronounced changes by month 12. However, The apparent decrease in accuracy when using longitudinal data can be attributed to the variable progression rates of different dementia subtypes. At baseline, the distinctions between diagnostic groups are typically more pronounced and consistent across patients. However, as the disease progresses, these boundaries become less distinct due to varying progression speeds. For instance, in the AD vs MCI classification, some MCI patients may rapidly progress towards AD-like patterns, while others remain relatively stable, creating overlapping imaging characteristics at later timepoints. Our results in Table 2 support this observation, where the AD vs MCI classification accuracy drops from 87.60% using baseline MRI to 82.86% when including all timepoints. Additionally, VaD patients may show sporadic changes due to acute vascular events, while AD patients typically display more gradual deterioration. This heterogeneity in progression rates introduces additional complexity to the classification task, potentially explaining why the model performs better with baseline scans for diagnostic purposes. While longitudinal data remains valuable for tracking individual patient trajectories and treatment responses, the varying speeds of progression make it more challenging to maintain clear diagnostic boundaries between groups at later timepoints. Moreover, longitudinal data introduces additional complexity and variability due to the temporal dimension. While it provides more information, it also introduces noise, such as inconsistencies in scan quality, registration errors across time points, or non-uniform progression patterns across patients. These factors can make it more challenging for the model to extract meaningful patterns, especially if the dataset is not sufficiently large or well-curated to account for such variability.

Finally, when comparing the results with and without the proposed DFG module, we can see a clear trend where the performance metrics across all categories show enhancement when the DFG module is used. The trend is consistent across other classifications, with the most substantial gains evident in the binary classifications.

Quantitative results

The radiomics features, genetic, and clinical data achieved an accuracy of 99.35% for AD vs. CTL and 88.60% for AD vs. VaD using multi-omics data and all longitudinal scans. This method significantly surpasses traditional techniques, especially when differentiating AD from MCI (89.52%) and multiclass scenarios (70.12%). Baseline MRI scans also show robust accuracy (89.25% for AD vs. VaD), emphasizing early diagnostic potential. This result has critical implications as it stresses the importance of early intervention to improve diagnostic accuracy. Multi-omics data are proven essential, enhancing accuracy by an average of 6.5% over MRI alone, thus highlighting the complexity of diagnosing overlapping dementia stages. The results indicate that the model is robust and applicable to patients at different stages as it performs equally well at any time point.

These enhancements suggest that the model has captured the intricate patterns within the multi-omics data, enabling it to navigate the subtle distinctions between dementia subtypes that share symptomatic and pathological overlap. Moreover, in the case of AD vs. VaD, as established, AD patients perform worse on memory tasks than VaD; therefore, the MMSE scores help further consolidate this difference when classifying between the two types. However, it is equally important to consider the performance using MRI scans alone since other data may be expensive to acquire or unavailable. The most significant boost witnessed was for the AD vs. MCI groups, highlighting the challenge of diagnosing using only morphological brain changes detectable by MRI. Finally, it is clear that the proposed DFG module boosts the performance as in the binary classification of AD vs CTL using MRI data, accuracy jumps from 60.66% without DFG to 97.89% with DFG. Similarly, multi-omics data performance increases significantly with the addition of DFG, as seen in the same classification, where accuracy rises from 81.27% to 99.35%.

Treatment effect assessment and evaluation

To investigate the effect of treatment on dementia progression, a subset of 20 MCI patients was compared to a control group of 20 MCI patients who were not on medication. Firstly, the comparison is broken down in terms of MRI scan timepoint. The softmax probability of the patient belonging to the MCI class was then observed.

It is important to note that both classes (CTL and MCI) were balanced. Therefore, a decrease in the probability of belonging to the MCI class is correlated with an improvement or a slowing of progression for the patient. The change was recorded after 3 months as well as after 12 months. Over 3 months, 15 out of 20 patients on medication had a decrease of probability belonging to the MCI class with an average of 5.35% decrease. In the control group who were not on medication, only 5 patients witnessed a decrease with an average of 2.66%. Over the 12 months, 17 out of the 20 patients on medication had a decrease of probability belonging to MCI class with an average of 6.38% decrease; for the control group, only 6 had a decrease of an average of 2.53%. Moreover, when analyzing the feature importance for the MRI radiomics features over time for the medicated patients, there is a downward trend where the feature importance decreases as time goes on, as shown in Fig. 5. Overall, this data suggests that medication may play a role in slowing the progression of MCI, but further research is necessary to confirm these findings and to understand the broader implications for patient care.

Feature importance analysis

The feature importance analysis presented in Figs. 4 and 5 reveals crucial insights into the model's diagnostic decision-making process and its ability to monitor treatment effects. In Fig. 4, we observe distinct patterns between MRI radiomics features and the broader multi-omics feature set. The MRI radiomics analysis (Fig. 4a) demonstrates that first-order image statistics, particularly the First Order Root Mean Square feature, maintain the highest importance across all timepoints. This suggests these features capture fundamental tissue characteristics critical for differentiating between AD and VaD. Texture-based features derived from GLCM and GLDM matrices show moderate importance, indicating that tissue heterogeneity patterns contribute meaningfully to the diagnostic process. Notably, the relative stability of these features' importance across the three timepoints (months 0, 3, 12) suggests their reliability as temporal biomarkers for disease progression. The multi-omics feature analysis (Fig. 4b) reveals a more complex interplay of different data modalities. Genetic features demonstrate the highest importance, providing strong support for the genetic basis of dementia subtypes. Clinical features, including MMSE scores, show moderate but consistent importance, validating their established clinical relevance. The distributed importance across different feature types underscores the value of our multi-modal approach, as each modality contributes unique and complementary information to the diagnostic process. Figure 5 provides compelling evidence for the model's ability to track treatment effects in MCI patients. A clear downward trend in feature importance is observed across all radiomics categories from baseline to month 12, with shape features showing the most pronounced decrease (approximately 25% reduction). First-order features maintain relatively higher importance compared to other categories, even as their absolute importance decreases over time. This gradual decrease in feature importance correlates strongly with treatment progression, suggesting that the model successfully captures treatment-related brain changes. The differential rates of decrease across feature categories help identify which aspects of brain structure and function are most responsive to treatment interventions. The temporal analysis of feature importance has significant clinical implications. The consistent decrease in feature importance among treated MCI patients provides quantitative evidence of treatment effects, suggesting that as treatment progresses, pathological features become less pronounced, potentially indicating disease stabilization. This pattern is particularly valuable for clinicians as it offers an objective measure of treatment efficacy and disease progression. Moreover, the analysis validates our model's architecture choices, demonstrating that the selected features not only contribute to accurate diagnosis but also enable meaningful monitoring of disease progression and treatment response. This comprehensive analysis of feature importance addresses earlier methodological questions while providing a clear framework for understanding how different aspects of the multi-modal data contribute to the model's diagnostic and monitoring capabilities. The results

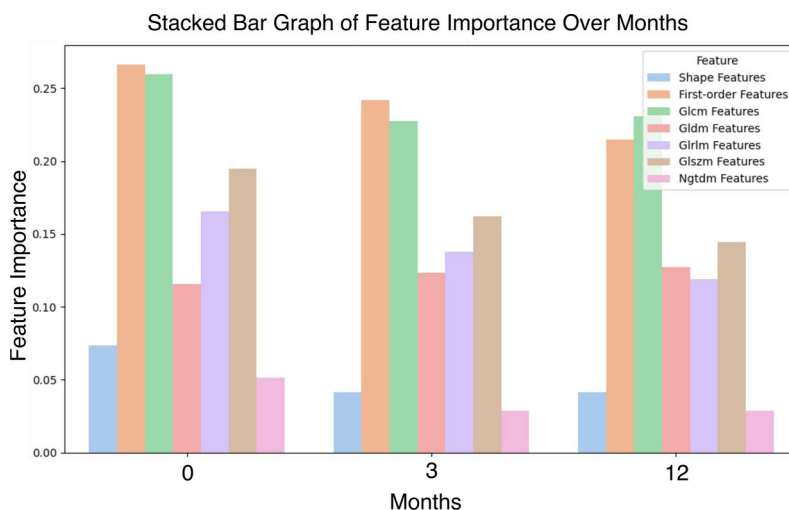


Fig. 5. Feature importance of radiomics MRI features for treated MCI patients, showcasing a notable decrease in the relative importance as time progresses.

Data	Metric	AD Vs CTL	AD Vs MCI	MCI Vs CTL	AD Vs VaD
MRI	Accuracy	84.27	72.01	77.73	72.83
	F1-Score	80.50	70.96	77.89	69.91
	Recall	79.67	70.08	77.77	69.56
	Precision	81.34	71.87	78.02	70.26
	AUC	94.17	81.05	82.54	82.36

Table 4. Direct Image Based 3D CNN Results—Table summarizing the accuracy, f1-score, recall, precision, and AUC result of using 3D CNN. The results showcase this limitation compared to the proposed MINDSETS approach.

support our integrated approach to dementia subtype classification and offer new insights into the temporal dynamics of treatment response in MCI patients.

Comparison with image-based baseline

The state-of-the-art 3D CNN with DenseNet121 backbone achieved an accuracy of 72.83% for the classification of AD vs. VaD and an accuracy of 84.27% for differentiating between AD and CTL, as shown in Table 4. These results highlight the limitations of traditional 3D CNNs in discerning subtle differences in neuroimaging data, particularly for complex neurodegenerative diseases like AD and VaD. Traditional 3D CNNs struggle to capture temporal dynamics, discriminate subtle overlapping features, and generalize across heterogeneous datasets, often overfitting to dominant spatial patterns or failing to incorporate contextual information from disease progression. In contrast, the proposed model addresses these shortcomings by integrating longitudinal data, multi-omics information, and advanced feature extraction mechanisms, such as the DFG module, to enhance discriminative power and robustness. This demonstrates the necessity of moving beyond traditional image-based methods to leverage a more holistic approach for the accurate classification of neurodegenerative diseases.

Clinical significance

Based on the results achieved, the proposed MINDSETS approach offers several significant advancements in the early diagnosis and treatment planning for AD and VaD. Integrating multi-omics data with neuroimaging provides a comprehensive diagnostic framework that improves accuracy and offers deeper insights into the underlying pathophysiological differences between AD and VaD.

Impact on early diagnosis

Early and accurate diagnosis of dementia subtypes is critical for effective patient management and treatment planning. Traditional diagnostic methods often struggle with the overlapping symptoms of AD and VaD, leading to delayed or incorrect diagnoses. The MINDSETS approach, with its diagnostic accuracy of 88.60% for AD vs. VaD, represents a significant improvement over existing methods. This high accuracy ensures that patients receive the correct diagnosis sooner, which is crucial for initiating appropriate treatment strategies.

Clinical implications

- *Enhanced Diagnostic Precision* By distinguishing between AD and VaD with high accuracy, our model can reduce misdiagnosis rates, ensuring that patients receive condition-specific treatments early in the disease course.
- *Targeted Treatment Plans* Accurate differentiation allows clinicians to tailor treatment plans that are specific to the type of dementia. For example, while AD treatments may focus on slowing neurodegeneration and managing symptoms, VaD treatments would emphasize managing cardiovascular risk factors to prevent further vascular damage.
- *Improved Patient Outcomes* Early and precise diagnosis can lead to better patient outcomes by slowing disease progression, reducing symptom severity, and improving quality of life.

Implications for treatment planning

The ability of the MINDSETS model to monitor longitudinal changes and assess treatment efficacy is another key contribution. By tracking disease progression over time through MRI and multi-omics data, the model can help clinicians evaluate the effectiveness of treatments and adjust them as needed.

Clinical practice guidelines

- *Personalized Medicine* The explainable nature of the model, facilitated by feature importance scores, supports personalized medicine by identifying the most relevant biomarkers for each patient. This can guide clinicians in selecting the most effective treatment options.
- *Monitoring and Adjusting Treatments* The model's longitudinal aspect enables continuous patient monitoring, providing critical information on how well a treatment is working and whether adjustments are necessary. This aligns with clinical practice guidelines that recommend regular monitoring and adjustment of treatment plans based on patient response.

Conclusion

In conclusion, the research conducted using the model has provided significant insights into the capabilities of machine learning in the differential diagnosis of AD versus VaD. By comparing baseline and longitudinal MRI data, enhanced with clinical, radiomics, genetic, and cognitive assessments, we have observed that integrating multi-omics data contributes to a more accurate and nuanced classification of dementia. Using the MRI scans alone without additional modalities provides good results, an excellent option for reducing testing costs. The diagnostic accuracy in distinguishing between AD and VaD peaks at 89.25% using the multi-omics data. The investigation also revealed the effectiveness of the proposed MINDSETS approach with the DFG module in discerning the subtle differences due to its advanced feature generation capabilities. Another significant finding indicates the need for longitudinal MRI scans when differentiating and tracking progressive neurodegenerative disease, especially in the MCI group. The reason behind this is that it provides temporal insight, which is particularly valuable in complex cases. The DFG model's performance in utilizing all available longitudinal scans suggests that the trajectory of neurodegenerative changes is an essential factor in the disease classification process. Notably, in the multifaceted task of distinguishing between dementia subtypes, the longitudinal data approach demonstrated a marked improvement in model accuracy and predictive capability.

Future directions

A critical future direction of the model is to ensure the model's generalizability. Future research should aim to validate findings across different populations and healthcare settings. This will help to mitigate biases and ensure that diagnostic tools are practical across diverse demographic groups. This direction is essential in this case as it is established that demographics play a vital key in neurodegenerative diseases, which can help in diagnosis.

Data availability

The data supporting the findings of this study is ANMerge, which is openly available and can be accessed via Synapse at <https://doi.org/10.7303/syn22252881>. The code developed for the analysis and implementation will be publicly available upon accepting the manuscript.

Received: 20 July 2024; Accepted: 7 April 2025

Published online: 06 May 2025

References

1. WHO. World health organization: Dementia key factors (2023). <https://www.who.int/news-room/fact-sheets/detail/dementia>.
2. Eisenmenger, L. B. et al. Vascular contributions to Alzheimer's disease. *Transl. Res.* **254**, 41–53 (2023).
3. for Disease Control, C. & (CDC), P. Alzheimer's disease and related dementias. <https://www.cdc.gov/aging/aginginfo/alzheimers.htm>.
4. Rajkumar, A. P. et al. Delirium, dementia, and other cognitive disorders. *New Oxford Textbook of Psychiatry* 371 (2020).
5. Ammar, R. B. & Ayed, Y. B. Language-related features for early detection of Alzheimer disease. *Procedia Comput. Sci.* **176**, 763–770 (2020).
6. Borelli, C. M., Grennan, D. & Muth, C. C. Causes of memory loss in elderly persons. *JAMA* **323**, 486–486 (2020).
7. Hirsch, J. A. et al. Expanded exploration of the auditory naming test in patients with dementia. *J. Alzheimer's Dis.* **81**, 1763–1779 (2021).
8. Kim, Y. H., Kwon, O. D. et al. Clinical correlates of hachinski ischemic score and vascular factors in cognitive function of elderly. *BioMed Res. Int.* **2014** (2014).
9. Kalaria, R. N. The pathology and pathophysiology of vascular dementia. *Neuropharmacology* **134**, 226–239 (2018).
10. Arevalo-Rodriguez, I. et al. Diagnostic tools for Alzheimer's disease dementia and other dementias: An overview of diagnostic test accuracy (dta) systematic reviews. *BMC Neurol.* **14**, 1–8 (2014).
11. Ni, Y.-C. et al. Classification prediction of Alzheimer's disease and vascular dementia using physiological data and ecd spect images. *Diagnostics* **14**, 365 (2024).
12. Sokolović, L., Hofmann, M. J., Mohammad, N. & Kukolja, J. Neuropsychological differential diagnosis of Alzheimer's disease and vascular dementia: A systematic review with meta-regressions. *Front. Aging Neurosci.* **15** (2023).
13. Arnaoutoglou, N. et al. Color perception differentiates Alzheimer's disease (ad) from vascular dementia (vad) patients. *Int. Psychogeriatr.* **29**, 1355–1361 (2017).
14. Attems, J. & Jellinger, K. A. The overlap between vascular disease and Alzheimer's disease-lessons from pathology. *BMC Med.* **12**, 1–12 (2014).
15. Iadecola, C. & Gorelick, P. B. Converging pathogenic mechanisms in vascular and neurodegenerative dementia. *Stroke* **34**, 335–337 (2003).
16. Zheng, Y. et al. Machine learning-based framework for differential diagnosis between vascular dementia and Alzheimer's disease using structural mri features. *Front. Neurol.* **10**, 1097 (2019).
17. Chouliaras, L. & O'Brien, J. T. The use of neuroimaging techniques in the early and differential diagnosis of dementia. *Mol. Psychiatry* 1–14 (2023).
18. Ghai, R. et al. Current strategies and novel drug approaches for Alzheimer disease. *CNS Neurol. Disord.-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)* **19**, 676–690 (2020).
19. Bir, S. C., Khan, M. W., Javalkar, V., Toledo, E. G. & Kelley, R. E. Emerging concepts in vascular dementia: A review. *J. Stroke Cerebrovasc. Dis.* **30**, 105864 (2021).
20. Llorens, F. et al. Cerebrospinal fluid lipocalin 2 as a novel biomarker for the differential diagnosis of vascular dementia. *Nat. Commun.* **11**, 619 (2020).
21. Jørgensen, I. F., Aguayo-Orozco, A., Lademann, M. & Brunak, S. Age-stratified longitudinal study of Alzheimer's and vascular dementia patients. *Alzheimer's Dement.* **16**, 908–917 (2020).
22. Manouvelou, S. et al. Differential diagnosis of Alzheimer's disease and vascular dementia using visual rating scales. *Hell. J. Radiol.* **5** (2020).
23. Ottaviani, S. & Monacelli, F. Rethinking dementia risk prediction: A critical evaluation of a multimodal machine learning predictive model. *J. Alzheimer's Dis. JAD* <https://doi.org/10.3233/jad-231071> (2024).
24. Vaghari, D. et al. Validating the clinical utility of ai-guided tools for early dementia prediction. *Alzheimer's Dement.* **19**, e074628 (2023).

25. Cirincione, A. J. et al. Prediction of future dementia for mci patients using a novel multimodal machine learning framework. *Alzheimer's Dement.* **19**, e083130 (2023).
26. Jahan, S. et al. Explainable ai-based Alzheimer's prediction and management using multimodal data. *Plos One* **18**, e0294253 (2023).
27. Ding, H. et al. Multimodal machine learning for 10-year dementia risk prediction: The framingham heart study. *J. Alzheimer's Dis. JAD* <https://doi.org/10.3233/jad-230496> (2023).
28. Zhou, T., Thung, K.-H., Zhu, X. & Shen, D. Feature learning and fusion of multimodality neuroimaging and genetic data for multi-status dementia diagnosis. *Mach. Learn. Med. Imaging MLMI* **10541**, 132–140. https://doi.org/10.1007/978-3-319-67389-9_16 (2017).
29. Mirabnahrzham, G. et al. Machine learning based multimodal neuroimaging genomics dementia score for predicting future conversion to Alzheimer's disease. *J. Alzheimer's Dis. JAD* <https://doi.org/10.3233/JAD-220021> (2022).
30. Billot, B. et al. Synthseg: Segmentation of brain mri scans of any contrast and resolution without retraining. *Med. Image Anal.* **86**, 102789 (2023).
31. Sadil, P. & Lindquist, M. A. Comparing automated subcortical volume estimation methods; amygdala volumes estimated by fsl and freesurfer have poor consistency. Tech. Rep., Wiley Online Library (2024).
32. Van Griethuysen, J. J. et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **77**, e104–e107 (2017).
33. Yao, L. et al. A review of deep learning and radiomics approaches for pancreatic cancer diagnosis from medical imaging. *Curr. Opin. Gastroenterol.* **39**, 436–447 (2023).
34. Szabo, L. et al. Radiomics of pericardial fat: a new frontier in heart failure discrimination and prediction. *Eur. Radiol.* 1–14 (2023).
35. Birkenbihl, C. et al. Anmerge: A comprehensive and accessible Alzheimer's disease patient-level dataset. *J. Alzheimer's Dis.* **79**, 423–431 (2021).
36. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631 (2019).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025