


# AI-driven multimodal precision diagnosis and progression prediction of Alzheimer's disease: Data fusion mechanisms, clinical applications, and research trends (2017–2024)

Wenhui Zhou<sup>1,\*</sup>, Yanhua Wang<sup>2,\*</sup>, Yudong Wu<sup>3,4</sup>, Xin Li<sup>4,5,6,7</sup>, Hong Liu<sup>1</sup>, Hailing Wang<sup>8</sup>, Zhichang Zhang<sup>2,†</sup> and He Huang<sup>1,†</sup> 

## Abstract

**Aims:** This study combines bibliometric and structured analyses to comprehensively examine the development, methodological characteristics, and application trends of multimodal artificial intelligence (AI) in Alzheimer's disease (AD) diagnosis.

**Materials and Methods:** Literature from January 1, 2017 to December 31, 2024, was retrieved from the Web of Science Core Collection. Retrospective bibliometric and visual analyses were conducted using VOSviewer, CiteSpace, and the Bibliometrix R package.

**Results:** A total of 234 papers were identified, showing a continuous increase in publication volume, with the United States and China as dominant contributors. The analysis focused on data modalities, fusion architectures, and clinical applications. Data trends highlight the fusion of imaging data with genetics, biomarkers, and clinical data. Methodologically, five fusion approaches were categorized, with intermediate fusion being the most widely used strategy for its ability to balance heterogeneous data integration. In application, multimodal AI demonstrated clear advantages in early diagnosis, disease classification, and progression prediction.

**Conclusion:** Research on multimodal AI for AD has gained global attention and remains a key direction for diagnostic innovation. By synthesizing bibliometric insights with structured analyses of modalities and fusion strategies, this study offers a systematic understanding of current progress and provides valuable guidance for future methodological and translational research.

## Keywords

Alzheimer's disease, multimodal artificial intelligence, early diagnosis, disease progression prediction, fusion network, biomarkers, cognitive decline

Received: 7 August 2025; accepted: 10 December 2025

<sup>1</sup>School of Medical Information Engineering, Shenyang Medical College, Shenyang, China

<sup>2</sup>Department of Computer, School of Intelligent Medicine, China Medical University, Shenyang, China

<sup>3</sup>School of Public Health, Shenyang Medical College, Shenyang, China

<sup>4</sup>Shenyang Key Laboratory of Prevention and Treatment of Systemic Important Diseases Associated with Oral Diseases, Shenyang, China

<sup>5</sup>School of Stomatology, Shenyang Medical College, Shenyang, China

<sup>6</sup>Key Laboratory of Human Ethnic Specificity and Phenomics of Critical Illness in Liaoning Province, Shenyang, China

<sup>7</sup>Key Laboratory of Phenomics in Shenyang, Shenyang, China

<sup>8</sup>Liaoning Jinqiu Hospital, Shenyang, China

\*These authors share the first authorship.

<sup>†</sup>These authors have contributed equally to this work.

## Corresponding authors:

Zhichang Zhang, Department of Computer, School of Intelligent Medicine, China Medical University, No. 77, Puhe Road, Shenbei New District, Shenyang City, Liaoning Province 110001, China.

Email: zczhang@cmu.edu.cn

He Huang, School of Medical Information Engineering, Shenyang Medical College, No. 146 Huanghe North Street, Yuhong District, Shenyang City, Liaoning Province 110034, China.

Email: syhuanghe@163.com



## Introduction

### Background

Alzheimer's disease (AD) is a common, irreversible neurodegenerative disorder and the leading cause of dementia among the elderly, posing a significant challenge to global public health. According to the latest epidemiological data, ~6.2 million people in the United States are currently living with AD, and it is projected that the number of cases worldwide will reach 152 million by 2050. With population aging accelerating, the burden imposed by AD and its associated cognitive decline (CD) has become increasingly pronounced, exerting substantial pressure on healthcare systems, patient families, and society at large.<sup>1</sup> At present, the treatment of AD primarily focuses on alleviating symptoms and improving the quality of daily life, with no effective cure available.<sup>2</sup> With the rapid advancement of artificial intelligence (AI) in the medical field, AI technology has emerged as a crucial research tool to enhance the accuracy of early diagnosis and treatment efficiency.<sup>3</sup>

Currently, the diagnosis of AD primarily relies on unimodal technologies, including imaging data, genetic data, biomarkers, and clinical data. Although these unimodal methods offer certain diagnostic value, they present notable limitations. For instance, the subtle pathological changes associated with early stage AD often exhibit low signal contrast in imaging, resulting in insufficient specificity.<sup>4</sup> Genetic testing, while useful for assessing disease risk, cannot provide a comprehensive prediction of an individual's likelihood of developing AD, and it may impose psychological stress and the risk of social discrimination on patients.<sup>5</sup> Cerebrospinal fluid (CSF), a key biomarker, requires an invasive collection process that increases patient discomfort and risks. Additionally, its diagnostic accuracy is constrained by sampling conditions and technical expertise.<sup>6</sup> Cognitive assessments can effectively evaluate cognitive impairment but are costly, time-consuming, and reliant on specialized personnel.<sup>7,8</sup> In contrast, multimodal data fusion leverages the complementary strengths of these various data sources, facilitating the construction of a more comprehensive disease profile. This integrative approach enhances the accuracy of early detection and diagnosis while providing more precise disease progression predictions.<sup>9,10</sup>

Multimodal AI technology plays a critical role in applying multimodal approaches. Machine learning models can automatically learn features from various modalities to construct robust predictive models. For example, Ran et al.<sup>11</sup> applied a multi-kernel regularized label relaxation linear regression model to integrate multimodal information, including magnetic resonance imaging (MRI), genetic data, and clinical cognitive behavioral test results, significantly improving diagnostic accuracy. Deep learning models excel at automatically extracting complex features from different modalities and integrating information at multiple

levels to uncover potential pathological changes and disease biomarkers. Convolutional neural networks (CNNs), for instance, are extensively used to analyze MRI and positron emission tomography (PET) imaging data, effectively capturing structural and functional alterations in the brain.<sup>12</sup> Recurrent neural networks (RNNs) are particularly suited for processing sequential data, enabling the analysis of temporal patterns in cognitive test results or physiological signals.<sup>13</sup> By combining imaging data, genetic data, clinical data, and biomarkers, multimodal AI enables comprehensive analysis and enhances the accuracy of early prediction and diagnosis of AD.<sup>14</sup>

In 2020, Lazli et al.<sup>15</sup> highlighted the limitations of using unimodal MRI for AD diagnosis, citing issues such as the large volume of data and low signal-to-noise ratio. They pointed out that integrating multiple imaging modalities, including structural MRI, functional MRI, CT, and PET, could enhance diagnostic accuracy and improve disease prediction capabilities. This study underscored the importance of multimodal fusion using imaging data for AD diagnosis, laying a solid foundation for future applications of multimodal data in Alzheimer's research. In 2021, Grueso et al.<sup>16</sup> summarized the application of multimodal data, including imaging data and clinical and genetic data, in AD, further promoting the application of multimodal data fusion and machine learning in the early diagnosis of AD. Compared to the 2020 research, it not only reviews the necessity of imaging data fusion, but also explicitly highlights the importance of integrating other modalities and deep learning models, thereby further improving diagnostic accuracy. In 2022, Qiu et al.<sup>17</sup> explored a fusion model combining MRI and clinical data, extending its application to other types of dementia. The study emphasized model interpretability, representing a significant advancement in multimodal data fusion and deep learning applications. Compared to the 2021 review, this research provided a more comprehensive perspective, offering greater insights into model explainability. In 2024, Teoh et al.<sup>18</sup> systematically reviewed various multimodal data fusion methods, including early, intermediate, and late. Unlike the 2022 review, the 2024 study provided a more structured and detailed analysis of specific fusion techniques, expanded the scope to cover a broader range of diseases, and emphasized addressing challenges related to data missingness, computational complexity, and model interpretability.

### Motivation for this article

Although several studies have summarized the progress of AI and multimodal data applications in AD research, there remains a lack of bibliometric analyses that quantitatively reveal the evolution of research hotspots and thematic trends in this field, as well as a lack of systematic examinations of multimodal AI development from three

complementary dimensions—data modalities, fusion mechanisms, and clinical applications.

To address these gaps, this study combines bibliometric analysis with a systematic structured analysis to investigate the development of multimodal AI in AD. The main contributions of this article are as follows:

1. Introducing a unified framework of five data fusion architectures. Building upon recent advances in multimodal learning, this study systematically categorizes and summarizes five representative fusion architectures—early fusion, intermediate fusion, late fusion, adaptive fusion, and hybrid fusion—that characterize how heterogeneous data are integrated in multimodal AI systems.

Each architecture represents a specific mechanism of information interaction: early fusion emphasizes direct feature-level integration; intermediate fusion combines learned representations; late fusion merges decisions at the output level; adaptive fusion dynamically adjusts strategies based on data reliability and task relevance; and hybrid fusion flexibly combines multiple fusion levels.

By systematically clarifying these architectures, this article establishes a more comprehensive perspective on how multimodal AI models process, integrate, and interpret diverse data sources in AD research.

2. Employing bibliometric and visualization analyses to reveal research evolution and hotspots. Using tools such as VOSviewer, CiteSpace, and Bibliometrix, this study quantitatively analyzes the literature from 2017 to 2024 to identify publication trends, collaboration networks, and thematic clusters. This bibliometric evidence enables the tracking of topic evolution and the prediction of emerging hotspots, providing a data-driven understanding of how multimodal AI has developed in the context of AD research.

By combining bibliometric mapping with literature analysis, this study bridges quantitative evidence with thematic insights, allowing a more comprehensive understanding of research evolution in multimodal AI for AD.

3. Establishing a three-dimensional (3D) analytical framework. To interpret the bibliometric findings and provide a holistic understanding, this article analyzes the multimodal AI literature across three complementary dimensions:

(1) Data modality dimension: Summarizes and compares the use of imaging, genetic, biomarker, and clinical data, revealing their complementarity and distinct diagnostic roles.

(2) Fusion architecture dimension: Summarizes the five representative data fusion architectures—early fusion, intermediate fusion, late fusion, adaptive fusion, and hybrid fusion—and outlines their core mechanisms and roles in multimodal integration.

(3) Clinical application dimension: Reviews the major application areas of multimodal AI in AD research, including early diagnosis, disease staging, progression prediction, biomarker identification, and other research.

Through this 3D analysis, the article outlines the relationships among data, fusion networks, and application domains, and reveals how multimodal AI research has evolved to address the heterogeneity of data sources and the methodological challenges in AD diagnosis, staging, and prognosis.

This integrated perspective, combining bibliometric analysis with structured analysis of data fusion methods and application domains, provides a comprehensive overview and useful reference for future research on multimodal AI in AD.

## Materials and methods

### Systematic search strategy

This study comprehensively analyzed existing literature in the Web of Science (WOS) database, focusing on the application of multimodal technologies in AD. A systematic search was performed on the WOS Core Collection, covering publications from 1 January 2017 to 31 December 2024. All relevant records were downloaded and completed within one day to minimize potential biases and accelerate the document retrieval process (10 January 2025).

The search strategy was developed using the following query:

TS = (("Artificial Intelligence" OR "Deep Learning" OR "Machine Learning" OR "Support Vector Machine" OR "Linear Regression" OR "Logistic Regression" OR "Decision Tree" OR "Random Forest" OR "K-Nearest Neighbors" OR "Naive Bayes" OR "Naive Bayes Model" OR "Convolutional Neural Network" OR "Recurrent Neural Network" OR "Fully Convolutional Network" OR "Generative Adversarial Network" OR "Reinforcement Learning" OR "Back Propagation" OR "Fully Neural Network" OR "Recursive Neural Network" OR "Autoencoder" OR "Deep Belief Network" OR "Restricted Boltzmann Machine" OR "Transformers" OR "Graph Convolution Networks" OR "K-Means" OR "AdaBoost" OR "Markov Chain" OR "Natural Language Processing" OR "Generative Pre-trained Transformer" OR "Bidirectional Encoder Representations from Transformers")AND ("Multimodal" OR "Fusion" OR "Data Fusion" OR "Combined Imaging" OR "Multi-Modality" OR "Multimodal Imaging Fusion" OR "Multimodal Learning")AND ("Alzheimer" OR "Alzheimer's Disease"))).

This search strategy ensured comprehensive coverage of AI models, multimodal fusion techniques, and AD research, providing a robust foundation for subsequent bibliometric analysis.

### Data inclusion criteria

The inclusion criteria for studies in the database were defined as follows: (1) published in English; (2) categorized

as original research articles or review articles (including systematic reviews); (3) published on or before 31 December 2024; and (4) focused on multimodal AI research in the context of AD. Studies meeting all these criteria were included in the analysis, while those failing to meet any were excluded, as illustrated in Figure 1.

A total of six papers were excluded for not being classified as articles or reviews, and 41 were excluded for not being indexed in SCI or SSCI. Additionally, 330 papers were excluded for being irrelevant to the research topic, based on the following predefined exclusion criteria:

(1) Non-multimodal studies: Excluded if they utilized only a single data modality (e.g. MRI alone, genetic data alone, or linguistic data alone) without explicit descriptions of “modality fusion,” “multimodal integration,” or equivalent operations. This also included studies applying multiple analytical methods to a single modality (e.g. using different algorithms to analyze the same MRI dataset).

(2) Non-core AD studies: Excluded if the primary research focus was not on AD, such as studies in which AD appeared only as an auxiliary or secondary validation example rather than the main subject of investigation.

(3) Non-AI-driven studies: Excluded if no AI methods—such as machine learning or deep learning—were used at any stage of the analytical workflow. That is, studies were excluded when all key processes (e.g. feature extraction, modality fusion, and decision modeling) were performed solely through manual rules or traditional statistical techniques without any learning-based algorithms.

To enhance the objectivity and consistency of screening, two researchers independently evaluated each article against the predefined criteria. Discrepancies were resolved through discussion with a third researcher, ensuring consensus. This process guaranteed the reproducibility and transparency of the screening workflow.

Ultimately, 234 records were extracted, each containing the abstract, title, authors, keywords, country/region, institution, journal, and references. The data from these 234 publications were exported as “plain text files” for bibliometric analysis using VOSviewer (version 1.6.19), CiteSpace (version 4.3.R1), and the bibliometrix package (version 4.3.0) in R (version 4.3.3). The study examined the following publication characteristics: countries/regions, journals, institutions, references, and keywords. Additionally, the h-index, a widely recognized metric for assessing the research productivity and academic impact of scholars,<sup>19</sup> was used to analyze both research productivity and academic influence in this study.

### Bibliometric data analysis

VOSviewer is a software tool that visualizes co-occurrence relationships among authors, keywords, citations, and other bibliometric elements. In VOSviewer visualizations, terms closer to each other are categorized into the same cluster. The size of a node represents the frequency of term

occurrence, while the distance between nodes indicates the strength of their association.<sup>20</sup> This study used VOSviewer (version 1.6.19) to cluster keywords for further analysis.

The bibliometrix package in RStudio enables comprehensive bibliometric analysis. By examining aspects such as the annual publication volume, the number of publications from various countries and institutions, and journal distribution, bibliometrix reveals key research hotspots, developmental trends, and leading contributors in multimodal technologies for AD. These insights provide a robust foundation for understanding the knowledge structure of this research domain and identifying future directions.

CiteSpace utilizes co-citation analysis and pathfinding algorithms to represent the evolution of a specific knowledge domain visually. Constructing a scientific knowledge map highlights relationships among publications, illustrating the historical research trajectory, current research landscape, and emerging topics in the field. CiteSpace offers valuable insights into the progression and prospects of research. This study applied CiteSpace for burst detection analysis of keywords to identify research hotspots and trending topics across different time periods, thereby facilitating a comprehensive analysis of research trends.<sup>21</sup>

## Result

### Annual publication volume analysis

Figure 2(a) shows that the annual publication volume on the application of multimodal AI in AD research has shown a significant upward trend since 2017. Between 2017 and 2019, the number of publications remained relatively low, indicating that the field was in its early stages. However, starting in 2020, the publication volume began to accelerate, with a notable increase in 2022 and 2023. This growth reflects the continuous development of multimodal AI technologies and their expanding applications in the AD domain. The increasing academic interest in this technology suggests that research output in this field is likely to grow further in the coming years, potentially driving breakthroughs in AD diagnosis, classification, disease progression prediction, and biomarker discovery.

### Analysis of national scientific output and collaboration networks

According to the visualization results of bibliometric data (Figure 3), research on multimodal AI in AD exhibits significant regional concentration globally. China and the United States dominate publication volume, with noticeably deeper geographic color markers than other countries, indicating the highest research output density in these two nations. South Korea, Spain, the United Kingdom, India,

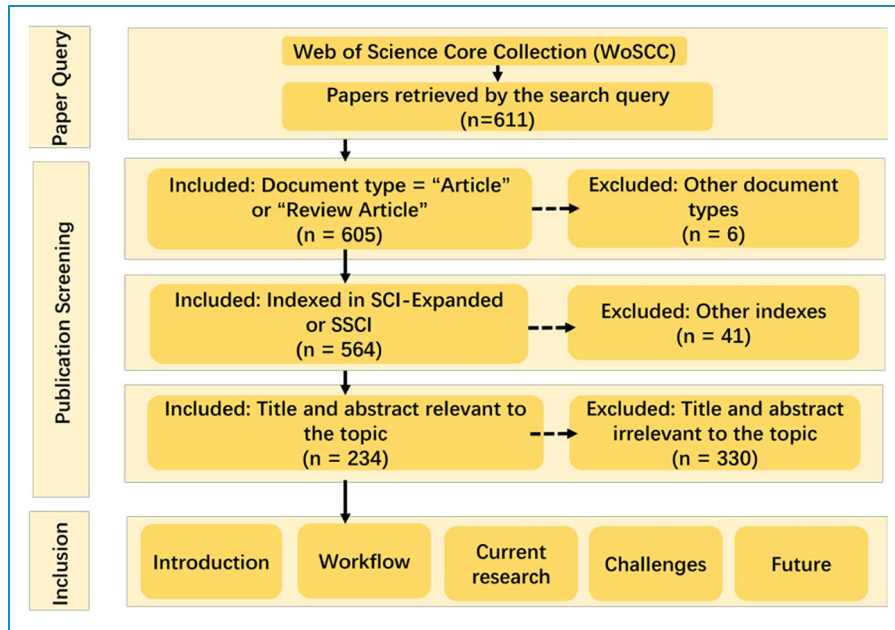


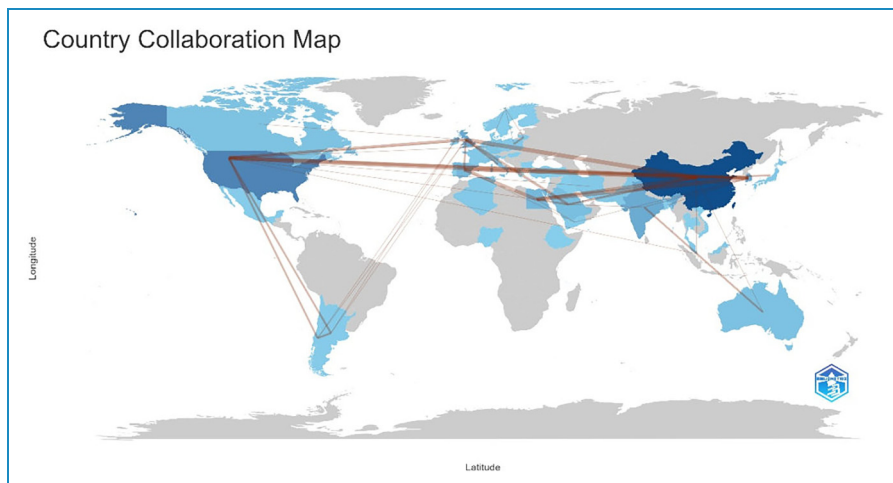
Figure 1. Criteria and flowchart for inclusion and exclusion.



Figure 2. Trends and characteristics of multimodal artificial intelligence (AI) research in Alzheimer’s disease from 2017 to 2024: (a) annual scientific production; (b) modal combination types; (c) fusion network architecture designs; and (d) clinical application scenarios.

and Egypt are closely followed, with gradually lighter color gradients, reflecting a sequential decrease in research productivity. Regarding international collaboration, China and the United States maintain relatively strong co-operation in this field, with a collaboration frequency of 10 instances. Additionally, China and the United Kingdom have

collaborated eight times, while South Korea and Egypt also show a collaboration frequency of eight instances. Similarly, Spain and Egypt have collaborated six times. These collaborative relationships illustrate global scientific communities’ interactive and cooperative efforts in advancing AD research through multimodal AI.



**Figure 3.** Distribution of publications by countries and regions.

### Journal analysis

Based on the data presented in Table 1, it is evident that the core journals in the field of AD research using multimodal AI exhibit a strong interdisciplinary nature. *Frontiers in Aging Neuroscience* leads with 13 publications (5.6% of the total) and has been cited 775 times (h-index = 7), highlighting its prominent role in neurodegenerative disease research. Its high publication and citation rates reflect the impact of the open-access model in promoting multidisciplinary research. In contrast, *NeuroImage* has published only seven papers (3.0% of the total) but accumulated 738 citations (h-index = 5), underscoring its significant influence in the intersection of brain imaging and AI. Neuroscience journals such as *Frontiers in Neuroscience* and the *Journal of Alzheimer's Disease* primarily focus on the clinical translation of multimodal data. Meanwhile, engineering and information science journals, including the *IEEE Journal of Biomedical and Health Informatics* and *Artificial Intelligence in Medicine*, emphasize algorithmic innovations in AI, representing the high academic barriers in technology-driven research. Notably, *Biomedical Signal Processing and Control* and *Computers in Biology and Medicine*, as leading journals in the biomedical engineering field, demonstrate a moderate publication volume and citation count. This reflects the practical orientation of signal processing technologies in facilitating the early diagnosis of AD.

### Institutional analysis

As shown in Table 2, the top 10 institutions have made significant academic contributions to the field of multimodal AI research on AD, as revealed by the bibliometric analysis. The Egyptian Knowledge Bank (EKB) and the Florida State University System lead in the number of publications. At

the same time, Shanghai Jiao Tong University has shown remarkable performance in citations and academic impact, underscoring its significant position in this field. With its substantial total citations, Boston University has emerged as one of the most academically influential institutions. Although Sungkyunkwan University (SKKU) and Fudan University have relatively fewer publications, they maintain strong citation records, indicating the sustained impact of their research. Overall, these institutions' research output and citation performance reflect their extensive participation and significant contributions to the advancement of multimodal AI research in AD, particularly in driving academic progress and enhancing the field's influence.

### Citation analysis

The analyzed papers in this study have been cited 7405 times, with the top 10 most cited references listed as shown in Table 3. The most cited paper, "Multimodal Classification of Alzheimer's Disease and Mild Cognitive Impairment,"<sup>22</sup> has received 60 citations. It proposed a multimodal classification approach integrating MRI, FDG-PET, and CSF biomarkers to enhance the early diagnosis of AD and mild cognitive impairment (MCI). An early fusion strategy was adopted to merge features from the three modalities, followed by classification using a kernel-combined multimodal method. Support vector machine (SVM) served as the classifier, with performance evaluated via 10-fold cross-validation. The results showed that the proposed method achieved a classification accuracy of 93.2% for distinguishing AD patients from healthy controls, outperforming the best single-modality result of 86.5%. For MCI versus healthy controls, the multimodal approach reached an accuracy of 76.4%, again significantly higher than single-modality methods. This study highlighted the advantages of multimodal data fusion in the early diagnosis of AD and MCI, particularly in enhancing sensitivity and

**Table 1.** Top 10 journals on artificial intelligence (AI)-based multimodal research in Alzheimer's disease (2017–2024).

Journal name	N	%	h	g	Cites	Q	IF
Frontiers in Aging Neurosci.	13	5.6	7	13	775	Q2	4.3
Biomed. Signal Process. Control	10	4.3	5	10	107	Q2	4.1
Frontiers in Neurosci.	8	3.4	5	8	237	Q2	5.1
NeuroImage	7	3.0	5	7	738	Q1	7.4
IEEE Access	7	3.0	4	7	204	Q3	3.4
J. Alzheimer's Dis.	7	3.0	3	5	25	Q1	4.0
Scientific Reports	7	3.0	3	7	472	Q2	3.9
Comput. Biol. Med.	6	2.6	5	6	198	Q2	4.6
IEEE J. Biomed. Health Inform.	6	2.6	5	6	493	Q1	6.6
Artif. Intell. Med.	6	2.6	4	6	75	Q1	7.5

**Table 2.** Top 10 contributing institutions in multimodal artificial intelligence (AI) research on Alzheimer's disease (2017–2024).

Affiliation	Articles	h	g	Cites
EGYPTIAN KNOWLEDGE BANK (EKB)	22	11	22	565
STATE UNIV. SYSTEM OF FLORIDA	22	10	15	257
SHANGHAI JIAO TONG UNIVERSITY	21	12	21	827
BOSTON UNIVERSITY	19	9	19	1559
SUNGKYUNKWAN UNIVERSITY (SKKU)	18	10	18	431
FUDAN UNIVERSITY	12	3	12	227
UNIVERSITY OF FLORIDA	11	6	10	106
CHINESE ACADEMY OF SCIENCES	10	5	10	1222
NATIONAL UNIVERSITY OF SINGAPORE	10	9	10	484
BENHA UNIVERSITY	9	7	9	326

**Table 3.** Top 10 most locally cited references in multimodal artificial intelligence (AI) research on Alzheimer's disease (2017–2024).

Title	Author(s)	Local citations
Multimodal classification of Alzheimer's disease and mild cognitive impairment	Daoqiang Zhang	60
The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods	Clifford R. Jack Jr.	37
Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis	Heung-Il Suk	37
Multimodal Neuroimaging Feature Learning for Multiclass Diagnosis of Alzheimer's Disease	Siqi Liu	35
Predicting Alzheimer's disease progression using multi-modal deep learning approach	Garam Lee	31
Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease	Daoqiang Zhang	31
2018 Alzheimer's disease facts and figures	Alzheimer's Association	29
Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis	Manhua Liu	28
A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease	Simeon Spasov	28
Random forest-based similarity measures for multi-modal classification of Alzheimer's disease	Katherine R. Gray	27

specificity. The proposed multimodal approach demonstrated excellent performance, illustrating that the combination of MRI, PET, and CSF biomarkers provides complementary information, offering a more precise diagnostic tool for early detection.



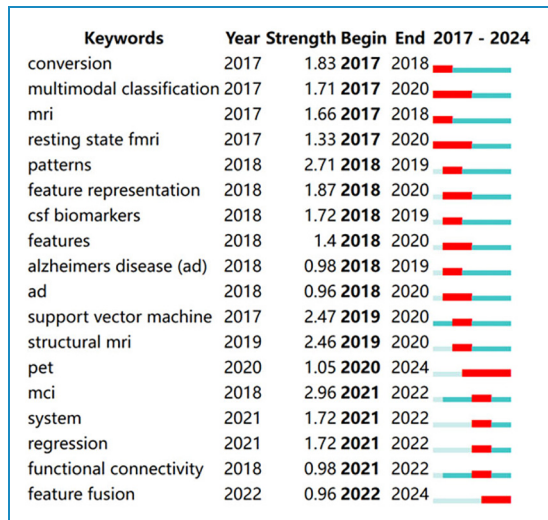


Figure 5. Top 18 keywords with the strongest citation bursts.

only focused on model performance but also on model interpretability and integrating multimodal data to enhance the diagnosis of AD and predict disease progression.

### Multimodal data characteristics

Multimodal data fusion has emerged as a crucial research direction, offering a more comprehensive understanding of disease mechanisms, predicting disease progression, and formulating personalized treatment plans by integrating different data types. For instance, combining genetic data with imaging data can reveal how genetic variations influence brain structure while integrating biomarkers with cognitive data can help assess early manifestations of the disease. This multimodal approach significantly enhances diagnostic accuracy and treatment effectiveness. The commonly used data modalities mainly include the following:

**Imaging data:** Common types of imaging data include MRI, PET, and diffusion tensor imaging (DTI). Imaging data provides detailed information on brain structure and function. PET, for instance, is instrumental in detecting metabolic changes in the brain, while conventional MRI is primarily used to reveal structural changes such as brain atrophy. Additionally, DTI is particularly useful for examining the integrity of white matter and assessing neural connectivity.

**Genetic data:** Genetic data typically include information on the APOE- $\epsilon 4$  gene and single-nucleotide polymorphisms (SNPs). Genetic data play a critical role in AD research. The APOE- $\epsilon 4$  gene is strongly associated with an increased risk of AD. SNPs provide more detailed genetic information, offering insights into the role of genetic factors in the onset and progression of AD.

**Biomarkers data:** Common biomarker data include amyloid-beta ( $A\beta$ ) and tau protein levels in CSF. These

are classical biomarkers of AD, reflecting the underlying biological processes of the disease. Measuring these biomarkers can facilitate early diagnosis, monitor disease progression, and evaluate treatment effectiveness.

**Clinical data:** Clinical data encompass demographic information such as name and age, as well as commonly used clinical assessments like cognitive test results, medication history, and comorbidities. These data provide valuable insights into the patient's health status, medical history, and treatment journey. Medication history and comorbidities may significantly influence the onset and progression of AD, offering critical background information for disease prediction and assessment.

**Other data:** Additional data types include retinal images and motor function data. Retinal imaging can aid in the early detection of neurodegenerative changes associated with AD, while motor function data reflect a patient's physical capabilities. Both types of data contribute valuable information for diagnosing and predicting AD.

### Clinical application scenarios

Multimodal AI in AD research can be applied to several key clinical scenarios:

**Early diagnosis:** Early diagnosis aims to detect signs of AD before symptoms become evident. By employing multimodal data fusion techniques, comprehensive analyses are conducted using medical imaging data (e.g. MRI and PET), genetic data (e.g. APOE- $\epsilon 4$  gene), clinical data (e.g. medical history and symptoms), and cognitive and behavioral assessment results (e.g. MMSE scores). Early diagnosis facilitates timely intervention and treatment, potentially delaying disease progression.

**Prediction of AD progression:** Multimodal data can be utilized to develop predictive models that estimate the rate and stage of AD progression. These models can forecast significant disease milestones by continuously monitoring changes in imaging data, cognitive function, and biomarker levels. This predictive capability supports clinicians in planning and adjusting treatment strategies in advance.

**Classification of AD stages:** Based on characteristic differences in multimodal data, AD can be classified into different stages, including MCI, early stage dementia, middle-stage dementia, and late-stage dementia. Accurate stage classification aids clinicians in assessing disease severity, providing a foundation for personalized treatment. It also enables researchers to conduct targeted studies focusing on specific stages of the disease.

**Diagnosis of AD:** Multimodal data integration, using machine learning or deep learning algorithms, enhances the accuracy and reliability of AD diagnosis. Disease-specific patterns can be identified by training models on large datasets comprising known AD patients and healthy controls. These models are then applied to new samples for precise diagnosis.

**Table 4.** Comparison of fusion strategies in multimodal artificial intelligence (AI) for Alzheimer's disease.

Fusion type	Fusion location	Interaction between modalities	Advantages	Disadvantages
Early fusion	Data input stage	Directly integrates multimodal data into a unified feature representation; interaction is simple or absent.	Easy to implement; suitable when modality correlation is strong; reduces model complexity.	Limited exploitation of complementary information; sensitive to missing/noisy data.
Intermediate fusion	After feature extraction, before decision layer	Features interact via attention mechanisms or shared layers in the middle of the model.	Captures inter-modality complementarity; enhances expression; fits weak correlation scenarios.	Higher complexity and training cost; needs good alignment between modalities.
Late fusion	Prediction (decision) layer	Modalities trained independently; results combined at decision level (e.g. voting and averaging).	Independent modality training; robust to missing data; suitable for heterogeneous modalities.	Weak inter-modality interaction; potential loss of complementary information.
Adaptive fusion	Throughout the model	Fusion strategies or weights are dynamically adjusted based on input data.	Adaptable to input characteristics; handles complex/missing modality cases well.	Complex model structure; difficult to train; requires large data volume.
Hybrid fusion	Multiple stages (input, middle, and output)	Combination of early, intermediate, and late fusion strategies.	Fully exploits complementary information; ideal for complex tasks; enhances performance.	High complexity; harder to optimize; fusion strategy design required.

**Biomarker research:** A primary research focus is the identification of biomarkers that reflect AD onset and progression, such as  $\beta$ -amyloid, phosphorylated tau protein, and specific genetic mutations. Multimodal data provide comprehensive insights from different perspectives, facilitating the discovery of disease-associated biomarkers. These biomarkers are essential for early diagnosis, disease monitoring, and drug development.

**Other research:** Additional research directions include pathological mechanism studies, personalized interventions, and cognitive function assessments.

**Pathological mechanism research:** Multimodal data are used to investigate the pathogenesis and pathological processes of AD from multiple dimensions, including genetic, cellular, tissue, and brain function levels.

**Personalized interventions:** Treatment and intervention plans are tailored based on the multimodal data characteristics of individual patients, promoting customized healthcare.

**Cognitive function assessment:** Multimodal data comprehensively evaluate patients' cognitive abilities, including memory, attention, language skills, and executive function. Combining cognitive test results, neuroimaging data, and physiological indicators provides a more accurate assessment of cognitive impairment and its progression, offering valuable information for diagnosis and treatment decisions.

## Fusion network architectures

This study involves five types of multimodal fusion architectures: early fusion, intermediate fusion, late fusion, adaptive fusion, and hybrid fusion. Only original studies that explicitly implemented multimodal fusion models were analyzed when classifying these architectures; review papers were not included in this methodological categorization. The structures and characteristics of these fusion methods are shown in Table 4.

### Early fusion

The architecture of early fusion is illustrated in Figure 6(a). In early fusion, multimodal data is integrated before the feature extraction layer, mapping heterogeneous data into a unified feature space for model processing. The core methods of early fusion include the following categories: (1) Concatenation<sup>23–35</sup>: This method directly concatenates data from different modalities along the feature dimension to generate a unified input vector. Its simplicity and efficiency make it advantageous while preserving the completeness of the original information. (2) Weighted averaging<sup>23,26,27,29,31,33,36,37</sup>: This approach assigns static weights to each modality to create a fused feature representation, highlighting the contributions of

key modalities. (3) Tensor fusion (Liu et al.,<sup>38</sup> Bi et al.,<sup>39</sup> Odujami et al.,<sup>40</sup> and Bi et al.<sup>41</sup>): Tensor fusion methods use high-order tensor operations (such as outer products and Kronecker products) to model fine-grained interactions between modalities. This strategy effectively captures non-linear interactions and enhances the representation capacity of the features.

From the perspective of data modalities, studies<sup>24–27,31,33,35,36,38,40,42–51</sup> have utilized imaging data (e.g. MRI and PET) for multimodal fusion. Among these, studies<sup>24–26,36,38,40,42–45,47–50</sup> used network-based approaches for fusion, employing CNNs and their variants, such as Inception-ResNet, ResNet-50, and mobile vision transformer. For example, ResNet-50, with its deep network architecture, can extract higher-level and more abstract features from imaging data, providing valuable insights for AD diagnosis and analysis. Some studies (Hojjati et al.,<sup>27</sup> Dachena et al.,<sup>29</sup> Gullett et al.,<sup>31</sup> Chen et al.,<sup>33</sup> Nan et al.,<sup>35</sup> Bi et al.,<sup>41</sup> Ismail et al.,<sup>46</sup> and Shukla et al.<sup>51</sup>) also applied network-based fusion but introduced innovations and adjustments in their network architecture and fusion methods based on their specific research objectives. For instance, some employed ensemble models to integrate multiple submodels, leveraging their advantages. Others used logistic regression (LR) models to further optimize classification and prediction based on fused data. Additionally, studies<sup>23,28–30,32,34,37,39,41,52–57</sup> fused imaging data with other modalities, such as biomarkers, clinical data, and cognitive and behavioral test data. For example, combining biomarker data (e.g. blood or CSF markers) with imaging data reflecting structural and functional brain changes can provide comprehensive insights into AD diagnosis and progression. The most commonly used network architectures in these studies include CNNs, RNNs, and their variants (e.g. gated recurrent units (GRUs), long short-term memory (LSTM)), SVMs, and RF. CNNs are effective for feature extraction from imaging data, RNNs excel in analyzing temporal patterns from behavioral and cognitive test data, and traditional machine learning algorithms like SVMs and RF are used for classification and prediction based on fused data. Some Jemimah et al.<sup>58</sup> and Zuo et al.<sup>59</sup> focused on modalities other than imaging data. For instance, Jemimah et al.<sup>58</sup> analyzed other types of data, such as motion data from depth cameras, force plates, and interface boards, to identify AD-related patterns. Zuo et al.<sup>59</sup> also utilized genetic data for model construction, investigating the genetic mechanisms and prediction methods for AD.

Regarding learning methods, studies<sup>27–31,33–35,37,39,41,51–55,57</sup> applied traditional machine learning algorithms, including SVMs, RF, decision trees DT, LR, and k-nearest neighbors (KNNs). Various feature selection techniques such as Fisher Score, greedy search, and genetic algorithms were used to optimize feature selection and enhance model performance. Multimodal time-series analysis, radiomics, and genetic-imaging data fusion were also commonly applied. Conversely, studies<sup>23–26,32,36,38,40,42–50,55,56</sup> employed deep

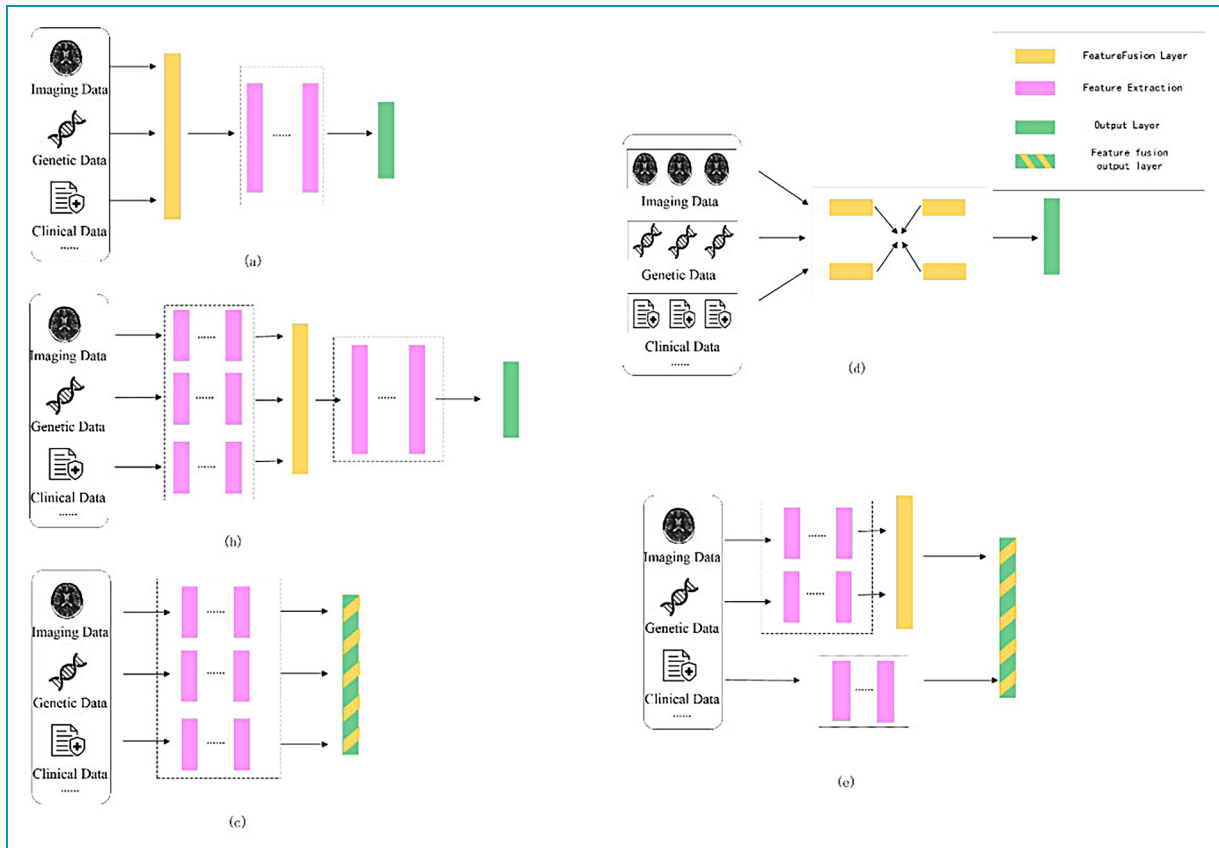
learning methods, including CNNs, RNNs, GRUs, deep belief networks (DBNs), and LSTMs. Advanced multimodal fusion techniques such as image fusion (MRI-PET), feature concatenation, and transfer learning were also adopted to improve model performance. Some studies further applied techniques like quantum dynamic optimization, wavelet transform, and ResNet architectures to enhance accuracy and robustness. Jemimah et al.<sup>58</sup> integrated machine learning with deep learning by combining a constrained deep learning model (c-Diadem) with KEGG pathway constraints while also using SHAP to interpret feature importance, offering novel approaches for AD research.

In terms of clinical applications, studies<sup>23–30,32–35,37–39,41,44–49,51–57</sup> focused on early diagnosis of AD, using multimodal data fusion and advanced learning methods to detect symptoms at an early stage for timely treatment and intervention. Studies<sup>23,27,28,30–32,34,36,37,41,44–46,55,56,58</sup> addressed the prediction of AD progression, aiming to predict disease progression trends through multimodal data analysis to support personalized treatment planning. Studies<sup>40,42,47,48,50,53,55</sup> focused on the classification of AD stages, dividing AD into different stages to aid physicians in tailoring treatment plans. Other studies<sup>30,38–40,42,43,48,50,55</sup> were dedicated to the diagnosis of AD, enhancing diagnostic accuracy through multimodal data and machine learning models. Furthermore, studies<sup>34,54,55,58</sup> focused on biomarker Research, identifying biomarkers indicative of AD progression to support diagnosis and monitoring. Bi et al.<sup>41</sup> and Jemimah et al.<sup>58</sup> contributed to pathological mechanism research, exploring the molecular and cellular mechanisms of AD to establish a theoretical basis for therapeutic development. Hall et al.<sup>57</sup> investigated personalized interventions, aiming to develop individualized treatment and intervention plans using multimodal data, thereby improving therapeutic outcomes and quality of life for patients.

Despite its advantages, early fusion still faces challenges when dealing with high-dimensional data, redundancy, and data quality issues. To enhance the effectiveness of early fusion, careful design in feature selection and dimensionality reduction is necessary to ensure that the fused data sufficiently represents the key information from each modality. In AD research and applications, early fusion has demonstrated significant value in supporting diagnosis and prediction, especially when integrating imaging data, clinical data, and biomarkers.

### Intermediate fusion

The architecture of intermediate fusion, as shown in Figure 6(b), involves merging data after feature extraction but before the decision layer. In this approach, different modalities interact and exchange information within the intermediate layers of the model. The main strategies for intermediate fusion include, (1) Joint representation learning methods<sup>59–63</sup>: These methods generate unified cross-modal representations through shared parameters or



**Figure 6.** Fusion network architectures for Alzheimer's disease multimodal data integration: (a) early fusion; (b) intermediate fusion; (c) late fusion; (d) adaptive fusion; and (e) hybrid fusion.

adversarial learning, creating low-dimensional, compact embeddings that reduce heterogeneity and enhance model robustness. (2) Cross-modal attention mechanisms<sup>64–69</sup>: By dynamically adjusting attention weights, these methods capture dependencies between modalities, adaptively focusing on key cross-modal relationships while preserving modality-specific information. (3) Multimodal recurrent network methods<sup>13,70–72</sup>: These methods apply temporal models to process dynamic multimodal data, fusing features within recurrent units. They are particularly effective for capturing disease progression trajectories in longitudinal studies. (4) Hybrid architecture fusion methods<sup>73–77</sup>: Combining heterogeneous networks such as CNNs and transformers, these approaches integrate multi-level features by leveraging complementary strengths, achieving more comprehensive feature fusion. These approaches effectively retain the individual characteristics of each modality while ensuring meaningful information exchange.

Based on the data modalities used in the studies, research works<sup>12,59–62,66,67,71,74,76,78–103</sup> utilized multimodal imaging data such as MRI, PET, and DTI as input, employing networks like 3D-CNN, ResNet, and vision transformer to extract high-order features. For instance, Zuo et al.<sup>59</sup> proposed the CT-GAN model, which integrates brain network

features from fMRI and DTI using a swapping bi-attention mechanism. Additionally, Lu et al.<sup>91</sup> utilized a multimodal and multiscale deep neural network to extract structural and functional features of the brain and their associations from MRI and FDG-PET imaging modalities. Studies<sup>13,14,63,64,68,69,72,73,75,77,79,99,104–143</sup> integrated imaging data with other modalities such as genetic, clinical, and biomarker data. Notable examples include the Qiang et al.<sup>104</sup> that constructed a dual attention network (DANMLP). This model used Patch-CNN to extract sMRI features, an MLP to encode APOE gene data, and positional and channel-wise self-attention for cross-modal alignment. El-Sappagh et al.<sup>70</sup> applied CNNs to process MRI data and BiLSTMs to model MMSE score sequences using multitask learning for AD progression prediction. In addition, Tabarestani et al.<sup>116</sup> proposed the kernelized tensorized multitask network (KTMnet) to simultaneously process MRI, PET, CSF, and cognitive scores, employing kernel mapping and tensor decomposition to achieve high-order feature interactions. Studies<sup>129,132,144–149</sup> explored non-imaging modalities such as biomarkers, speech, and eye-tracking data. For instance, Ilias and Askounis<sup>149</sup> a method that combines BERT for text processing and vision transformer (ViT) for analyzing speech spectrograms,

capturing the associations between linguistic and acoustic features through co-attention. Yin et al.<sup>144</sup> designs an Internet of Things (IoT)-based architecture that utilizes eye-tracking nodes and machine learning algorithms on the cloud platform to achieve early screening of AD.

Based on the applied learning methods, studies<sup>14,83–86,88,97,99,100,102,107,108,112–114,120,122,123,125,130,135–137,140,142,144</sup> employed traditional machine learning approaches. For instance, Castellazzi et al.,<sup>83</sup> Kim and Lee,<sup>107</sup> and Mehdipour Ghazi et al.<sup>112</sup> utilized shallow models such as SVM, RF, and extreme learning machines (ELMs). Specifically, Kim and Lee,<sup>107</sup> proposed a multi-scale hierarchical extreme learning machine (MSH-ELM) that integrates sMRI, FDG-PET, and CSF features. AlMohimeed et al.<sup>140</sup> applied particle swarm optimization (PSO) to select cognitive sub-scores and improved classification performance using a multi-level stacked ensemble model.

On the other hand, studies<sup>12–14</sup> employed deep learning methods, including CNN, RNN, GRU, DBN, and LSTM networks. These studies also applied multimodal fusion techniques such as image fusion (e.g. MRI-PET), feature concatenation, and transfer learning to enhance the models' capability in processing multimodal data. Optimization techniques like quantum dynamic optimization, wavelet transform, and ResNet architecture were introduced to improve model performance and accuracy. Morar et al.<sup>13</sup> proposed a highly accurate predictive model by integrating multimodal data and deep learning techniques, which can effectively predict cognitive test scores for AD patients. Miao et al.<sup>74</sup> combine machine learning and deep learning to propose a multimodal multiscale transformer fusion network for computer-aided diagnosis of AD. This model not only efficiently extracts local abnormal information related to AD but also learns the underlying representations of multimodal data, thereby significantly improving diagnostic accuracy.

In terms of clinical applications, studies<sup>64,65,67–69,72,76–79,81,82,87–91,93,95–97,100,101,104–109,111,112,115,120,121,123–129,132–134,136–145,147–149</sup> focused on early diagnosis of AD using intermediate fusion methods. By integrating various data sources and fusion strategies, these studies aimed to detect potential pathological features at an early stage, providing a scientific basis for early intervention and treatment. Studies<sup>12–14,68,70,72,75,77,79,83,94,102,108,109,112–116,118,120,122,127,131,134,138,140,148</sup>

focused on AD progression prediction, employing multimodal data and appropriate models to capture dynamic changes during disease development. The goal was to support the development of personalized treatment plans and disease management strategies. Research works<sup>12,61,63,66,67,76,78,80–86,90,91,93,95,97,98,101,102,112,117,123,130,132,135–137,139,142,143,146</sup>

were dedicated to AD stage classification. By analyzing and extracting features of different disease stages, these studies accurately classified various stages of AD, providing deeper insights into disease progression. In the field of AD diagnosis,

studies<sup>63,69,71,73,74,94,96,101,110,117,119,122,128,129,133,137,146,147</sup> adopted comprehensive approaches that combined multiple data sources and fusion strategies. This facilitated improved diagnostic accuracy and reliability. Furthermore, studies<sup>13,64,73,95,105,107,111–113,116,120,122,127,134,135</sup> concentrated on biomarker research. By integrating and analyzing multimodal data, these studies aimed to identify biomarkers associated with AD, providing novel biological targets for diagnosis, treatment, and further research.

Intermediate fusion methods utilize multi-level and dynamic feature interaction mechanisms, effectively preserving the unique characteristics of each modality (e.g. structural information from MRI and metabolic features from PET) while enabling in-depth exploration of cross-modal semantic relationships. Compared to early fusion, which involves simple feature concatenation, and late fusion, which relies on independent decisions, intermediate fusion achieves a more refined balance at the feature representation level. This approach provides a robust and interpretable solution for multimodal diagnosis of AD.

### Late fusion

The architecture of late fusion is illustrated in Figure 6(c). Late fusion is a strategy that integrates multimodal information during the prediction phase of a model. Its core concept is to independently process each modality's data and combine the results using flexible decision-making strategies. Compared to early and intermediate fusion, late fusion effectively preserves modality-specific information and reduces the complexity of aligning heterogeneous data. Based on different fusion strategies, late fusion can be categorized into the following approaches: (1) Voting mechanism<sup>150–153</sup>: This approach aggregates predictions from multiple unimodal models using majority voting or weighted voting strategies to generate the final decision. By independently training classifiers for each modality, it maximizes the retention of unique modality features and mitigates the impact of noise or misclassification from a single modality, thereby enhancing decision robustness. (2) Multiple kernel learning (MKL)<sup>11,154</sup>: In this method, multiple kernel functions are constructed to learn features from different modalities. By performing feature fusion in various kernel spaces, MKL improves the overall classification or prediction performance. (3) Stacking ensemble<sup>155</sup>: Utilizing the ensemble integration (EI) framework, local predictive models are trained separately from each modality of data, and the complementary and consensus information from multimodal data are integrated to predict the likelihood of patients with MCI progressing to dementia in the future.

In terms of the data modalities used in the studies, research works<sup>11,150,154,156–159</sup> primarily focused on multimodal fusion using imaging data, integrating sMRI, fMRI, and PET imaging, or incorporating MEG to enhance

spatiotemporal resolution. Common fusion strategies include feature-level fusion (e.g. kernel canonical correlation analysis (KCCA) and fractal dimension analysis) and decision-level fusion (e.g. majority voting and multi-kernel SVM), with an emphasis on capturing the complementary information between imaging modalities. For example, Wang et al.<sup>156</sup> proposed a fusion strategy using KCCA to integrate sMRI and fMRI features. On the other hand, studies<sup>151,152,155,160–165</sup> combined imaging data with non-imaging modalities, such as biomarkers (e.g. CSF), behavioral scales (e.g. MMSE), genetic information (e.g. ApoE4), or sociodemographic data, to enhance the clinical interpretability of models. For instance, Zhang et al.<sup>163</sup> introduced an FCRN network to extract imaging features from MRI/PET, which were then fused with clinical data processed using an MLP at the decision level. Zhang et al.<sup>160</sup> constructed a graph neural network (GNN) to combine sMRI/PET imaging features with phenotypic data, achieving personalized early diagnosis of AD. Meanwhile, studies<sup>153,166,167</sup> explored non-imaging approaches, relying primarily on physiological signals (e.g. EEG and HRV), behavioral data (e.g. movement and speech), and biomarkers, focusing on the temporal dynamics or behavioral patterns. These studies can be broadly classified into spatiotemporal feature fusion and biomarker-driven approaches. In the spatiotemporal feature fusion category, Boudaya et al.<sup>151</sup> successfully implemented a hybrid machine learning model to integrate the time–frequency features of EEG and HRV, achieving a classification accuracy of 93.86% for MCI. Mohan Gowda et al.<sup>153</sup> combined a long-term recurrent convolutional network (LRCN) with a RF classifier to fuse video and sensor data for fall detection.

In terms of learning methods, studies<sup>11,152,155,157,159,165,168</sup> primarily applied machine learning techniques, including ensemble learning, SVM, and traditional regression models, which are widely used in small-sample and heterogeneous data scenarios. For example, Cirincione et al.<sup>155</sup> employed an EI framework to fuse MRI, PET, and clinical data for predicting the conversion of MCI, achieving an AUC of 0.81. Similarly, in the context of deep learning, studies<sup>150,151,153,154,156,158,160,161,163,165,169</sup> adopted advanced methods such as 3D CNN, GNNs, and federated learning to enhance the modeling of complex modalities. For instance, Zhang et al.<sup>160</sup> constructed a multimodal GNN that integrated graph structures from sMRI/PET with phenotypic information. Lakhan et al.<sup>165</sup> proposed the FDCNN-AS framework, utilizing federated learning to enable cross-institutional multimodal data fusion while ensuring data privacy. Additionally, Cheng et al.<sup>161</sup> converted one-dimensional fNIRS signals into two-dimensional Gramian angular fields (GAFs) and applied CNNs for end-to-end feature learning.

In clinical applications, studies<sup>11,150,154,155,157–165</sup> were primarily focused on the early diagnosis of AD, leveraging the complementary nature of imaging data and biomarkers.

For example, El-Gamal et al.<sup>158</sup> combined structural MRI and 11C PiB PET for regional diagnosis, achieving a sensitivity of 96.77%. In the context of AD progression prediction, studies<sup>152–154,157,161,166,170</sup> emphasized dynamic functional connectivity and the temporal changes of biomarkers. Chen et al.<sup>159</sup> demonstrated the predictive value of both static and dynamic features from resting-state fMRI for monitoring MCI conversion. Alty et al.<sup>166</sup> developed a smartphone application called TapTalk, which assesses AD risk by analyzing hand movement and speech patterns; regarding AD stage classification, studies<sup>11,151,156</sup> integrated neuroimaging and cognitive assessments. For instance, Wang et al.<sup>156</sup> applied KCCA to fuse sMRI and fMRI features for AD staging. Studies<sup>150,162,167</sup> were employed to diagnose AD, while studies<sup>152</sup> specifically investigated biomarkers. Additionally, Mohan Gowda et al.<sup>153</sup> applied multimodal techniques in the caregiving domain, introducing a multimodal fall detection system that combined visual and sensor data to enhance caregiving efficiency.

By integrating independent modeling with flexible decision-making, late fusion preserves modality-specific characteristics while balancing inter-modality complementarity. The voting mechanism is widely applied for its simplicity and efficiency, multi-kernel learning demonstrates advantages in handling heterogeneous data, and stacking ensemble methods leverage meta-learning to achieve high-level feature interactions. These approaches have shown significant advantages in improving the accuracy of AD diagnosis, advancing the understanding of pathological mechanisms, and optimizing caregiving practices. Future research should further explore integrating federated learning and explainability-enhancement techniques to accelerate the translation of multimodal AI into clinical diagnostic applications.

### Adaptive fusion

The architecture of adaptive fusion is shown in Figure 6(d). Adaptive fusion allows models to achieve optimal integration based on information from different modalities. Its core concept involves adaptive weight adjustment or domain adaptation to optimize the fusion process. It enables the model to automatically learn which modalities or features are critical for the final prediction. Adaptive fusion strategies can be categorized into the following approaches: (1) Dynamic weight adjustment<sup>168–171</sup>: Methods such as attention mechanisms,<sup>168,171</sup> Adaboost,<sup>172</sup> and boosting ensemble approaches<sup>169</sup> are used to dynamically allocate modality weights, optimizing feature contributions. This strategy offers considerable flexibility and robustness. (2) Cross-modal generation and imputation<sup>172–181</sup>: GANs and RNN are employed to generate missing data dynamically, enhancing the completeness of multimodal data. This method effectively addresses the problem of modality

missingness in clinical data collection, expanding the applicability of models. (3) Self-supervised learning and mutual information maximization (Fedorov et al.<sup>182</sup>): Contrastive learning or maximizing mutual information can be applied to discover cross-modal consistency without requiring labeled data. This method enhances the model's generalization ability and improves cross-modal consistency in learning.

From the perspective of the data modalities used, studies<sup>168,170,171,176–184</sup> applied multimodal fusion using only imaging data, with a particular emphasis on addressing data missingness through GANs in studies.<sup>172–174,180,181</sup> These approaches effectively resolve data gaps by generating missing modalities. Additionally, studies<sup>168,171,176</sup> utilized attention mechanisms to dynamically allocate weights for imaging data fusion, enhancing model interpretability and performance. Studies<sup>169,180,181,183,185</sup> integrated imaging data with other modalities, such as biomarkers and clinical data. Specifically, Zhang et al.<sup>169</sup> and Hwang et al.<sup>180</sup> combined imaging and clinical data using feature augmentation strategies, employing various dimensional models or generative models to strengthen feature fusion. In contrast, Nguyen Martí-Juan et al.<sup>181</sup> integrates imaging data, clinical data, and time-series data, leveraging the minimal RNN to accommodate missing data and to learn the relationships between modalities across different time points. Moreover, the Chandrashekar et al.<sup>186</sup> did not use imaging data but instead leveraged genetic data and other modalities. It introduced the DeepGAMI model, which utilizes functional genomic interaction networks to connect data from different sources. It effectively integrated these modalities for comprehensive analysis through feature imputation and adaptive fusion.

In terms of learning methods, studies<sup>168,171–177,183–189</sup> primarily applied deep learning approaches, utilizing architectures such as RNNs, CNNs, GANs, self-supervised learning, and deep generative models for multimodal data fusion. These methods emphasize end-to-end learning, enabling the automatic extraction of high-level cross-modal features while reducing reliance on manual feature engineering. As demonstrated in studies,<sup>168,171,172</sup> and shared latent spaces, as in Martí-Juan et al.<sup>183</sup> study, attention mechanisms were used to adaptively allocate the importance of different modalities. Meanwhile, studies<sup>169,170,178</sup> integrated machine learning and deep learning techniques, employing deep learning models for feature extraction and generation and applying machine learning classifiers for decision-making. This hybrid approach formed an ensemble classification network, enhancing overall model robustness and accuracy.

In the context of clinical applications, studies<sup>151,153,155–157,162,164,165,169,170</sup> focused on early diagnosis of AD, employing adaptive fusion strategies to effectively integrate data from different modalities. This approach significantly supports the early identification of AD, facilitating timely

interventions. Regarding the classification of AD stages, studies<sup>168,170,172,182,186</sup> applied adaptive fusion techniques to combine multiple modalities, achieving more accurate differentiation between various stages of AD. This enables the development of more targeted treatment plans. In the field of prediction of AD progression, studies<sup>177,180,181,183,186</sup> utilized adaptive fusion strategies to dynamically capture the temporal relationships among multimodal data. This enhanced the accuracy of predicting disease progression, providing valuable insights for disease management. For the diagnosis of AD, studies<sup>170,171,173,175,178,179,185</sup> employed adaptive fusion methods, which facilitated comprehensive information extraction from different modalities. This improved the accuracy and reliability of AD diagnosis. In the domain of biomarker research, the Liu et al.<sup>175</sup> integrated multimodal data to identify potential biomarkers associated with AD. This approach offers new therapeutic targets and advances understanding the disease's underlying mechanisms.

Adaptive fusion optimizes the multimodal data fusion process by automatically adjusting the weights of different modalities or using generative models to fill in missing data. This strategy demonstrates significant flexibility and robustness, enabling the model to autonomously learn the most suitable fusion approach based on the characteristics of the data. Through methods such as dynamic weight adjustment, cross-modal generation and imputation, joint learning, and multi-task optimization, adaptive fusion has shown remarkable application potential in AD research. It exhibits exceptional adaptability and practicality in scenarios involving missing or heterogeneous data.

### Hybrid fusion

The architecture of hybrid fusion is illustrated in Figure 6(e). Hybrid fusion leverages the advantages of early, intermediate, and late fusion strategies to achieve complementary and synergistic multimodal data integration at different levels. The core architecture, as illustrated in the figure, primarily includes the following four modes: (1) Early + Intermediate fusion<sup>187,188</sup>: This approach integrates raw data during the low-level feature extraction phase (early fusion) while enhancing feature representation through cross-modal interactions at the intermediate level (intermediate fusion). By retaining fine-grained information from the original modalities and capturing complex dependencies through dynamic intermediate-layer interactions, this method significantly enhances model robustness. (2) Early + Late fusion<sup>190</sup>: In this approach, low-level features are fused early to preserve detailed information, followed by integrating multimodal classification results at the decision layer (late fusion). This strategy effectively balances local feature preservation with global decision optimization, offering strong biological interpretability. (3) Intermediate + Late fusion<sup>191</sup>: This method generates joint feature representations at the intermediate layer, while the late fusion

stage employs generative models to address modality missingness. The key advantage of this approach lies in its end-to-end optimization of both the generation and classification processes, thereby enhancing clinical applicability. (4) Comprehensive hybrid fusion<sup>192</sup>: Characterized by its application of multi-stage strategies combining early, intermediate, and late fusion, this approach is well-suited for complex data scenarios. It flexibly handles heterogeneous data while ensuring feature complementarity and model generalization capability.

From the perspective of the data modalities used in the studies, approaches employing imaging and non-imaging multimodal data<sup>187,188,189,190,191</sup> integrated MRI/PET with clinical, genetic, and cognitive score data. For instance, Rahim et al.<sup>187</sup> integrated 3D MRI imaging data, cognitive scores, and biomarker modalities using a hybrid architecture comprising 3D CNNs and bidirectional RNNs (BRNNs) to efficiently predict the progression of AD. In contrast, non-imaging multimodal studies<sup>192</sup> relied on longitudinal clinical data (e.g. MMSE and RAVLT) and demographic features. Through LSTM-based temporal modeling and ensemble learning, these approaches optimized predictive stability.

From the perspective of learning methods, studies<sup>188,190</sup> integrated machine learning and deep learning approaches, using artificial neural networks (ANNs) or SVM to enhance feature interpretability while employing CNN to process high-dimensional imaging data, thereby balancing performance and clinical applicability. For instance, Tu et al.<sup>188</sup> expanded clinical features using geometric algebra, fused them with CNN-derived imaging features, and then input the results into an ANN for classification. Studies<sup>187,191–193</sup> applied deep learning methods, utilizing end-to-end architectures such as transformers and LSTM-CNNs to automate feature extraction and fusion. Notably, Nguyen et al.<sup>181</sup> used a 3D-CNN combined with a BRNN to integrate longitudinal MRI data with baseline biomarkers, achieving an AUC of 0.96.

In clinical applications, hybrid fusion strategies have been widely utilized in various aspects of AD research. Studies<sup>187,188,194,195</sup> focused on the early diagnosis of AD. By integrating multimodal data through hybrid fusion, these studies provided robust support for the early detection of AD, facilitating timely intervention and treatment during the initial stages of the disease. Studies<sup>187,190–192</sup> applied hybrid fusion to predict AD progression. Leveraging this strategy, these studies captured disease progression trends more accurately, offering valuable insights for developing personalized treatment plans and disease management strategies. Balaji et al.<sup>193</sup> addressed the classification of AD stages. Fusing multimodal data using hybrid fusion achieved the precise classification of different AD stages, aiding clinicians in better understanding the disease's progression and formulating targeted treatment plans.

Hybrid fusion leverages the advantages of early, intermediate, and late fusion strategies to achieve complementary and synergistic integration of multimodal data across multiple levels. This approach effectively preserves the detailed features of each modality while enhancing model robustness and performance through fusion at different stages, thereby improving the model's adaptability to complex data patterns.

To better understand the distribution of fusion strategies and their associated learning methods, Table 5 categorizes the reviewed studies according to fusion type (early, intermediate, late, adaptive, and hybrid) and whether they employ machine learning, deep learning, or a combination of both.

## Discussion

### Summary of current research status

This section provides an in-depth analysis of the current state of multimodal AI applications in AD research from three perspectives: data modalities used, fusion network architectures, and clinical application scenarios. Only original research articles that proposed or implemented multimodal AI models were included in this analytical section; review papers were not considered in these methodological and application-oriented analyses.

The characteristics of multimodal data are illustrated in Figure 2(b); integrating imaging data with other data types has increasingly become a mainstream approach. Specifically, the combination of multiple imaging modalities remains predominant, highlighting the significant role of imaging data in AD research. However, studies integrating imaging data with other data modalities, such as biomarkers, clinical data, and genetic data, are more prevalent. This reflects a growing trend towards multimodal fusion, which enables a more comprehensive extraction of disease-related information, thereby enhancing diagnostic accuracy and disease prediction capabilities. In contrast, the combination of non-imaging data remains relatively limited, potentially due to factors such as data heterogeneity and insufficient sample sizes. Overall, future research is expected to focus more on optimizing fusion strategies that integrate imaging and non-imaging data. For instance, combining imaging modalities with biomarker data could further improve model performance and enhance its clinical applicability. The design of fusion network architectures is illustrated in Figure 2(c); intermediate fusion is the most widely applied fusion network architecture, primarily because it effectively captures cross-modal semantic associations while preserving the distinct characteristics of each modality. This approach provides more robust and interpretable solutions for the diagnosis and prediction of AD. Although early fusion offers advantages such as streamlined processes, it faces challenges in handling high-

**Table 5.** Distribution of learning methods applied to different data fusion strategies.

Fusion method	Machinelearning	Deeplearning	Machine learning + deep learning	Articles
Early fusion	✓			[See references <sup>7,28,29,30,31,33,34,35,37,39,41,51,52,53,54,57</sup> ]
		✓		[See references <sup>23,24,25,26,32,36,38,40,42,43,44,45,46,47,48,49,50,55,56</sup> ]
Intermediate fusion	✓			[See references <sup>14,83,84,85,86,88,97,99,100,102,107,108,112,113,114,120,122,123,125,130,135,136,137,140,142</sup> ]
		✓		[See references <sup>13,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,77,79,80,87,89,90,91,12,92,93,94,95,96,98,101,103,104,105,109,110,115,116,117,118,121,124,126,127,129,131,132,133,134,138,139,141,143,147,149,184</sup> ]
			✓	[See references <sup>7,6,78,81,82,106,111,119,128,137,145,146</sup> ]
Late fusion	✓			[See references <sup>11,1152,155,157,159,162,164,166</sup> ]
		✓		[See references <sup>150,151,153,154,156,158,160,161,163,165,167,189,196</sup> ]
Adaptive fusion		✓		[See references <sup>168,171,173,174,175,172,176,177,179,180,181,186,182,183,185</sup> ]
			✓	[See references <sup>170,169,178,197</sup> ]
Hybrid fusion		✓		[See references <sup>187,191,192,193</sup> ]
			✓	[See references <sup>188,190,198</sup> ]

dimensional data, information redundancy, and data quality issues. Late fusion, on the other hand, focuses on integrating information at the decision-making stage, resulting in limited cross-modal interaction. Adaptive fusion and hybrid fusion demonstrate significant potential due to their capacity for dynamic integration and enhanced performance. However, their relatively limited application can be attributed to the high model complexity. With advancements in computational power and algorithm optimization, these fusion strategies are expected to gain broader adoption in the future.

Clinical application scenarios are illustrated in Figure 2(d); early diagnosis has garnered the most attention, with numerous studies focusing on detecting signs of AD before symptoms become apparent. This emphasis reflects the critical importance of early intervention in improving patient outcomes. MCI represents a key stage in the early diagnosis of AD, while CD is typically regarded as an earlier prodromal manifestation that precedes MCI.<sup>194</sup> As such, CD is implicitly encompassed within the AD progression continuum rather than analyzed as a separate category in our framework. In many of the included studies, multimodal AI models involved participants with early CD or MCI, allowing these models to capture preclinical cognitive changes that occur prior to a formal AD diagnosis. Accordingly, the integrated use of multimodal AI can effectively identify cognitive alterations that precede the clinical onset of AD, demonstrating its relevance across the continuum from early CD to overt AD. Beyond early diagnosis, research on AD progression prediction and stage classification is also substantial, contributing to the development of personalized treatment plans and enhancing therapeutic effectiveness. While studies on biomarkers are relatively fewer in number, they hold significant value, as identifying more effective biomarkers is crucial for early diagnosis, disease monitoring, and drug development. Other research areas, such as pathological mechanism exploration and personalized interventions, remain limited in the number of studies. However, they are of great importance for gaining a deeper understanding of AD pathogenesis and enhancing the precision of treatment strategies.

### Limitations and challenges of existing research

Although multimodal AI has made significant progress in AD research, several limitations and challenges remain:

In terms of multimodal data characteristics, issues related to data quality and consistency are particularly prominent. Multimodal data from different sources often exhibit variations in acquisition equipment, processing methods, and annotation standards, leading to increased noise and bias that can compromise model performance and generalization capability. For instance, MRI devices from different hospitals may have varying parameters, resulting in inconsistent image quality, which hinders the

model's ability to accurately extract features during the learning process. Additionally, the lack of sufficient data poses a significant challenge, especially for rare diseases or specific AD subtypes where data scarcity limits model training and validation. This data insufficiency prevents the model from fully capturing the complex characteristics and underlying patterns of the disease.

Each fusion approach has limitations from the perspective of fusion network architectures. Early fusion requires extensive data preprocessing and struggles to handle high-dimensional data and information redundancy. Intermediate fusion, while capable of capturing cross-modal interactions, often involves high model complexity, prolonged training time, and strict requirements for modality alignment. Late fusion retains modality-specific information but lacks sufficient cross-modal interactions, limiting its ability to fully leverage complementary information. Although adaptive and hybrid fusion methods offer notable advantages, their complex model structures pose significant training challenges, requiring substantial computational resources and domain expertise. Additionally, in practical applications, the interpretability of these models remains a concern, making it difficult to intuitively understand the rationale behind model decisions.

Furthermore, the clinical translation of multimodal AI in AD research faces significant challenges. Most current studies remain at the laboratory stage, with a considerable gap between research findings and practical clinical applications. Clinicians often have limited understanding and acceptance of complex AI models, and the reliability and safety of these models require further validation in real-world clinical settings. For instance, the model's performance may vary across different ethnicities and age groups. Ensuring that the model functions reliably and accurately across diverse clinical scenarios remains a critical issue that needs to be addressed.

At the same time, ethical concerns cannot be overlooked. The collection and utilization of multimodal data involve issues related to patient privacy protection and data security. For example, the use of genetic data may pose privacy risks, and improper access or misuse of such data could lead to severe social and psychological consequences for patients. Additionally, AI models may exhibit decision-making biases. Ensuring model fairness and impartiality to prevent adverse impacts on specific demographic groups is a critical issue that requires careful consideration.

### Future research directions

Future research will focus on further optimizing multimodal fusion strategies, particularly the integration of imaging and non-imaging data. Non-imaging modalities—such as biomarkers data, clinical data, genetic data—offer complementary information that can enrich imaging-based analysis. Their integration is expected to strengthen AI applications across key clinical domains of AD research, including early diagnosis, prediction of disease progression, classification

of disease stages, clinical diagnosis, and biomarker discovery. Although current studies are most concentrated in early diagnosis, expanding multimodal integration to these diverse application areas will help establish a more comprehensive understanding of AD pathophysiology and improve translational potential. With technological advances, the seamless combination of heterogeneous data will enable more accurate and generalizable predictive models, providing strong support for disease monitoring, treatment optimization, and personalized intervention frameworks. In terms of fusion network architecture design, although intermediate fusion currently dominates research, adaptive fusion, and hybrid fusion strategies are likely to become prominent research focuses in the future, driven by advances in computational power and algorithm optimization. These methods offer greater flexibility and robustness, enabling dynamic adjustment of fusion strategies based on data characteristics and making them particularly suitable for handling heterogeneous data and addressing data missingness issues. Furthermore, with the continuous development of deep learning and self-supervised learning techniques, future research should also prioritize enhancing model interpretability and practicality to better support clinical decision-making and implementation. From the perspective of clinical applications, multimodal AI will continue to drive innovations across multiple domains of AD research. Current studies are predominantly concentrated on early diagnosis, which remains the most active and impactful application area, followed by prediction of disease progression, classification of AD stages, and clinical diagnosis, reflecting growing efforts to enhance disease stratification and monitoring accuracy. In addition, biomarker-driven research—though relatively limited in number—has demonstrated strong potential for elucidating disease mechanisms and identifying therapeutic targets, thereby supporting individualized treatment optimization. By integrating diverse data sources, including imaging, biomarker, clinical, and genetic information, multimodal AI systems can provide more comprehensive assessments of disease trajectories and enable tailored intervention strategies. Notably, recent research<sup>195</sup> has confirmed that single-modal AI has achieved robust performance in early CD diagnosis, while also highlighting the potential of multimodal integration—such as combining MRI with clinical or speech-based data—to further enhance diagnostic efficacy in clinical settings. Building on these findings, and considering the close clinical and pathological continuum between CD and AD, future research should extend the application of multimodal AI from AD to the broader CD spectrum, thereby improving its generalizability and translational value across related cognitive disorders.

Overall, future research on multimodal AI in AD should prioritize the refinement of data fusion methodologies, the optimization of fusion architectures, and the advancement of clinical translation. To ensure reproducibility and clinical reliability, efforts should focus on standardization sample imbalance and missing modalities. In addition, enhancing model


interpretability and transparency will be essential for establishing clinician trust and facilitating regulatory acceptance. By bridging methodological innovation with clinical applicability, multimodal AI has the potential to provide robust tools for early diagnosis, disease monitoring, and personalized treatment in AD and related cognitive disorders.

## Analysis and conclusion

This study conducted a bibliometric and visual analysis of the applications of multimodal AI in AD. Through the analysis of existing literature, it was observed that multimodal AI technologies based on imaging modalities demonstrate significant application value in early diagnosis, disease prediction, and stage classification. Emerging fusion strategies, such as intermediate fusion and adaptive fusion, have proven more effective in handling complex multimodal data, thereby enhancing model performance and robustness.

Despite the remarkable progress made by multimodal AI in AD research, challenges remain, including data heterogeneity, computational complexity, and model interpretability. Future research should focus on optimizing multimodal data fusion strategies, improving model interpretability, and leveraging generative models and adaptive learning methods to address issues related to data scarcity and noise. With continuous technological advancements, multimodal AI is expected to play an increasingly pivotal role in early diagnosis, disease progression prediction, and personalized interventions for AD. This will further drive the development of precision medicine and provide more scientifically informed decision support for clinical practice.

## ORCID iD

He Huang  <https://orcid.org/0009-0005-9166-1350>

## Ethical approval

No patients were involved in this research, so the ethical approval is not required. The data sets analyzed during this study are available in the Web of Science core collections.

## Author contributions

Wenhui Zhou: conceptualization, methodology, software, and writing—original draft. Yanhua Wang: software and writing—review. Yudong Wu: visualization and investigation. Xin Li: supervision. Hong Liu: software. Hailing Wang: validation. Zhichang Zhang: conceptualization and validation. He Huang: writing—review and editing.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the International Industrial Technology Research and Development Project of the Liaoning Provincial Department of Science and Technology (No. 2025JH2/101900023);

the Liaoning Provincial Department of Education (Nos. LJ222410164024 and LJ222410164030); the Liaoning Province Science and Technology Joint Plan (Technology Research and Development Program Project, No. 2024JH2/102600237); the Shenyang Medical College Horizontal Research Projects (Nos. SYKT2025002, SYKT2025005, and SYKT2025007); and the Liaoning Provincial Department of Education Project “Research and Exploration on Olfactory Training Methods and Devices for Anti-Drug Squirrels” (Grant No. SYYX201909).

### Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### References

- Vitali F, Branigan GL and Brinton RD. Preventing Alzheimer’s disease within reach by 2025: targeted-risk-AD-prevention (trap) strategy. *Alzheimer’s Dement: Transl Res Clin Interv* 2021; 7: e12190.
- Nordberg A, Rinne JO and Kadir A, et al. The use of PET in Alzheimer disease. *Nat Rev Neurol* 2010; 6: 78–87.
- Zhang YD, Dong Z and Wang SH, et al. Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation. *Information Fusion* 2020; 64: 149–187.
- Jack CR, Wiste HJ and Weigand SD, et al. Age, sex, and apoe  $\epsilon$ 4 effects on memory, brain structure, and  $\beta$ -amyloid across the adult life span. *JAMA Neurol* 2015; 72: 511–519.
- Karlawish J. Addressing the ethical, policy, and social challenges of preclinical Alzheimer disease. *Neurology* 2011; 77: 1487.
- Shaw LM, Arias J and Blennow K, et al. Appropriate use criteria for lumbar puncture and cerebrospinal fluid testing in the diagnosis of Alzheimer’s disease. *Alzheimer’s Dement J Alzheimer’s Assoc* 2018; 14: 1505–1521.
- Muller S, Herde L and Preische O, et al. Diagnostic value of digital clock drawing test in comparison with cerad neuropsychological battery total score for discrimination of patients in the early course of Alzheimer’s disease from healthy individuals. *Sci Rep* 2019; 9: 3543.
- Olson R, Parkinson M and McKenzie M. Selection bias introduced by neuropsychological assessments. *Canad J Neurol Sci* 2010; 37: 264–268.
- Oduami M, Maskeliunas R and Damaeviius R, et al. Machine learning with multimodal neuroimaging data to classify stages of Alzheimer’s disease: a systematic review and meta-analysis. *Cogn Neurodyn* 2024; 18: 20.
- Liu X, Chen K and Wu T, et al. Use of multimodality imaging and artificial intelligence for diagnosis and prognosis of early stages of Alzheimer’s disease. *Transl Res* 2018; 194: 56–57.
- Ran X, Shi J and Chen Y, et al. Multimodal neuroimage data fusion based on multikernel learning in personalized medicine. *Front Pharmacol* 2022; 13: 947657.
- Gravina M, García-Pedrero A and Gonzalo-Martín C, et al. Multi input–multi output 3D CNN for dementia severity assessment with incomplete multimodal data. *Artif Intell Med* 2024; 149: 102774.
- Morar U, Martín H and Izquierdo W, et al. Prediction of cognitive test scores from variable length multimodal data in Alzheimer’s disease. *Cognit Comput* 2023; 15: 2062–2086.
- Minhas S, Khanum A and Riaz F, et al. Predicting progression from mild cognitive impairment to Alzheimer’s disease using autoregressive modelling of longitudinal and multimodal biomarkers. *IEEE J Biomed Health Inform* 2017; 22: 818–825.
- Lazli L, Boukadoum M and Mohamed OA. A survey on computer-aided diagnosis of brain disorders through MRI based on machine learning and data mining methodologies with an emphasis on Alzheimer disease diagnosis and the contribution of the multimodal fusion. *Appl Sci* 2020; 10: 1894.
- Grueso S and Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer’s disease dementia: a systematic review. *Alzheimers Res Ther* 2021; 13: 1–29.
- Qiu S, Miller MI and Joshi PS, et al. Multimodal deep learning for Alzheimer’s disease dementia assessment. *Nat Commun* 2022; 13: 3404.
- Teoh JR, Dong J and Zuo X, et al. Advancing healthcare through multimodal data fusion: a comprehensive review of techniques and applications. *PeerJ Comput Sci* 2024; 10: e2298.
- Cabezas-Clavijo A and Delgado-López-Cózar E. Google scholar and the h-index in biomedicine: the popularization of bibliometric assessment. *Med Intensiva (English Edition)* 2013; 37: 343–354.
- Van Eck N and Waltman L. Software survey: Vosviewer, a computer program for bibliometric mapping. *Scientometrics* 2010; 84: 523–538.
- Chen C and Song M. Visualizing a field of research: a methodology of systematic scientometric reviews. *PLoS ONE* 2019; 14: e0223994.
- Zhang D, Wang Y and Zhou L, et al. Multimodal classification of Alzheimer’s disease and mild cognitive impairment. *Neuroimage* 2011; 55: 856–867.
- Shetty J, Shetty NP and Kothikar H, et al. Detection of Alzheimer’s disease progression using integrated deep learning approaches. *Intell Automat Soft Comput* 2023; 37: 1345–1362.
- Kang L, Jiang J and Huang J, et al. Identifying early mild cognitive impairment by multi-modality MRI-based deep learning. *Front Aging Neurosci* 2020; 12: 206.
- Huang Y, Xu J and Zhou Y, et al. Diagnosis of Alzheimer’s disease via multi-modality 3D convolutional neural network. *Front Neurosci* 2019; 13: 509.

26. Wisely CE, Wang D and Henao R, et al. Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging. *British J Ophthalmol* 2022; 106: 388–395.
27. Hojjati SH, Ebrahimzadeh A and Khazaei A, et al. Predicting conversion from MCI to AD by integrating RS-FMRI and structural MRI. *Comput Biol Med* 2018; 102: 30–39.
28. El-Sappagh S, Saleh H and Sahal R, et al. Alzheimer's disease progression detection model based on an early fusion of cost-effective multimodal data. *Future Generat Comput Syst* 2021; 115: 680–699.
29. Dachena C, Casu S and Fanti A, et al. Combined use of MRI, FMRI and cognitive data for Alzheimer's disease: preliminary results. *Appl Sci* 2019; 9: 3156.
30. Dong N, Fu C and Li R, et al. Machine learning decomposition of the anatomy of neuropsychological deficit in Alzheimer's disease and mild cognitive impairment. *Front Aging Neurosci* 2022; 14: 854733.
31. Gullett JM, Albizu A and Fang R, et al. Baseline neuroimaging predicts decline to dementia from amnesic mild cognitive impairment. *Front Aging Neurosci* 2021; 13: 758298.
32. Alqahtani N, Alam S and Aqeel I, et al. Deep belief networks (DBN) with iot-based Alzheimer's disease detection and classification. *Appl Sci* 2023; 13: 7833.
33. Chen Z, Bi S and Shan Y, et al. Multiparametric hippocampal signatures for early diagnosis of Alzheimer's disease using 18f-fdg PET/MRI radiomics. *CNS Neurosci Ther* 2024; 30: e14539.
34. KP MN and Thiagarajan P. Feature selection using efficient fusion of Fisher score and greedy searching for Alzheimer's classification. *J King Saud Univ-Comput Inform Sci* 2022; 34: 4993–5006.
35. Nan F, Li S and Wang J, et al. A multi-classification accessment framework for reproducible evaluation of multimodal learning in Alzheimer's disease. *IEEE/ACM Trans Comput Biol Bioinform* 2022; 21: 559–572.
36. Sharma R, Goel T and Tanveer M, et al. Conv-ervfl: convolutional neural network based ensemble RVFL classifier for Alzheimer's disease diagnosis. *IEEE J Biomed Health Inform* 2022; 27: 4995–5003.
37. Minhas S, Khanum A and Alvi A, et al. Early MCI-to-AD conversion prediction using future value forecasting of multimodal features. *Comput Intell Neurosci* 2021; 2021: 6628036.
38. Liu F, Yuan S and Li W, et al. Patch-based deep multimodal learning framework for Alzheimer's disease diagnosis using multi-view neuroimaging. *Biomed Signal Process Control* 2023; 80: 104400.
39. Bi X, Cai R and Wang Y, et al. Effective diagnosis of Alzheimer's disease via multimodal fusion analysis framework. *Front Genet* 2019; 10: 976.
40. Odusami M, Damaševičius R and Milieškaitė-Belousovienė E, et al. Alzheimer's disease stage recognition from MRI and PET imaging data using pareto-optimal quantum dynamic optimization. *Heliyon* 2024; 10: e34402.
41. Bi X, Li L and Xu R, et al. Pathogenic factors identification of brain imaging and gene in late mild cognitive impairment. *Interdiscip Sci: Comput Life Sci* 2021; 13: 511–520.
42. Fu J, Li W and Du J, et al. Multimodal medical image fusion via Laplacian pyramid and convolutional neural network reconstruction with local gradient energy strategy. *Comput Biol Med* 2020; 126: 104048.
43. Rallabandi VS and Seetharaman K. Deep learning-based classification of healthy aging controls, mild cognitive impairment and Alzheimer's disease using fusion of MRI-PET imaging. *Biomed Signal Process Control* 2023; 80: 104312.
44. Fakoya AA and Parkinson S. A novel image casting and fusion for identifying individuals at risk of Alzheimer's disease using MRI and PET imaging. *IEEE Access* 2024; 12: 134101–134114.
45. Zhang X, Lin W and Xiao M, et al. Multimodal 2.5D convolutional neural network for diagnosis of Alzheimer's disease with magnetic resonance imaging and positron emission tomography. *Progress In Electromag Res* 2021; 171: 21–34.
46. Ismail WN, PP FR and Ali MA. A meta-heuristic multi-objective optimization method for Alzheimer's disease detection based on multi-modal data. *Mathematics* 2023; 11: 957.
47. Dwivedi S, Goel T and Tanveer M, et al. Multimodal fusion-based deep learning network for effective diagnosis of Alzheimer's disease. *IEEE MultiMedia* 2022; 29: 45–55.
48. Dai Y, Bai W and Tang Z, et al. Computer-aided diagnosis of Alzheimer's disease via deep learning models and radiomics method. *Appl Sci* 2021; 11: 8104.
49. Odusami M, Maskeliūnas R and Damaševičius R, et al. Explainable deep-learning-based diagnosis of Alzheimer's disease using multimodal input fusion of PET and MRI images. *J Med Biol Eng* 2023; 43: 291–302.
50. Kong Z, Zhang M and Zhu W, et al. Multi-modal data Alzheimer's disease detection based on 3D convolution. *Biomed Signal Process Control* 2022; 75: 103565.
51. Shukla A, Tiwari R and Tiwari S. Analyzing subcortical structures in Alzheimer's disease using ensemble learning. *Biomed Signal Process Control* 2024; 87: 105407.
52. Sánchez-Reyna AG, Celaya-Padilla JM and Galvan-Tejada CE, et al. Multimodal early Alzheimer's detection, a genetic algorithm approach with support vector machines. In: *Healthcare*, Vol. 9, p.971. MDPI.
53. Jahan S, Abu Taher K and Kaiser MS, et al. Explainable AI-based Alzheimer's prediction and management using multimodal data. *PLoS ONE* 2023; 18: e0294253.
54. Wang Y and Wang X. Exploration of imaging genetic biomarkers of Alzheimer's disease based on a machine learning method. *J Integr Neurosci* 2024; 23: 81.
55. El-Sappagh S, Saleh H and Ali F, et al. Two-stage deep learning model for Alzheimer's disease detection and

- prediction of the mild cognitive impairment time. *Neural Comput Appl* 2022; 34: 14487–14509.
56. Chang CY, Slowiejko D and Win N. Prediction and clustering of Alzheimer's disease by race and sex: a multi-head deep-learning approach to analyze irregular and heterogeneous data. *Sci Rep* 2024; 14: 26668.
  57. Hall JB, Akter S and Rao P, et al. Feasibility of using a novel, multimodal motor function assessment platform with machine learning to identify individuals with mild cognitive impairment. *Alzheimer Disease Assoc Disord* 2024; 38: 344–350.
  58. Jemimah S, AlShehhi A and Initiative ADN. c-diadem: a constrained dual-input deep learning model to identify novel biomarkers in Alzheimer's disease. *BMC Med Genomics* 2023; 16: 244.
  59. Zuo Q, Shen Y and Zhong N, et al. Alzheimer's disease prediction via brain structural-functional deep fusing network. *IEEE Trans Neural Syst Rehabil Eng* 2023; 31: 4601–4612.
  60. Tu Y, Lin S and Qiao J, et al. Multimodal fusion diagnosis of Alzheimer's disease based on FDG-PET generation. *Biomed Signal Process Control* 2024; 89: 105709.
  61. Cao Y, Kuai H and Liang P, et al. Bnloop-gan: a multi-loop generative adversarial model on brain network learning to classify Alzheimer's disease. *Front Neurosci* 2023; 17: 1202382.
  62. Odusami M, Maskeliūnas R and Damaševičius R. Pareto optimized adaptive learning with transposed convolution for image fusion Alzheimer's disease classification. *Brain Sci* 2023; 13: 1045.
  63. Tong T, Gray K and Gao Q, et al. Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. *Pattern Recognit* 2017; 63: 171–181.
  64. Kim SK, Duong QA and Gahm JK. Multimodal 3D deep learning for early diagnosis of Alzheimer's disease. *IEEE Access* 2024; 12: 46278–46289.
  65. Cai H, Gao Y and Liu M. Graph transformer geometric learning of brain networks using multimodal MR images for brain age estimation. *IEEE Trans Med Imaging* 2022; 42: 456–466.
  66. Tang C, Wei M and Sun J, et al. Csagp: detecting Alzheimer's disease from multimodal images via dual-transformer with cross-attention and graph pooling. *J King Saud Univ-Comput Inform Sci* 2023; 35: 101618.
  67. Tang C, Xi M and Sun J, et al. Macfnct: detection of Alzheimer's disease via multiscale attention and cross-enhancement fusion network. *Comput Methods Programs Biomed* 2024; 254: 108259.
  68. Machado Reyes D, Chao H and Hahn J, et al. Identifying progression-specific Alzheimer's subtypes using multi-modal transformer. *J Pers Med* 2024; 14: 421.
  69. Golovanevsky M, Eickhoff C and Singh R. Multimodal attention-based deep learning for Alzheimer's disease diagnosis. *J Am Med Inform Assoc* 2022; 29: 2014–2022.
  70. El-Sappagh S, Abuhmed T and Islam SR, et al. Multimodal multitask deep learning model for Alzheimer's disease progression detection based on time series data. *Neurocomputing* 2020; 412: 197–215.
  71. Feng C, Elazab A and Yang P, et al. Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBI-LSTM. *IEEE Access* 2019; 7: 63605–63618.
  72. Rahim N, Abuhmed T and Mirjalili S, et al. Time-series visual explainability for Alzheimer's disease progression detection for smart healthcare. *Alexandria Eng J* 2023; 82: 484–502.
  73. Kadri R, Bouaziz B and Tmar M, et al. Efficient multimodal method based on transformers and coatnet for Alzheimer's diagnosis. *Digit Signal Process* 2023; 143: 104229.
  74. Miao S, Xu Q and Li W, et al. Mmtfn: multi-modal multi-scale transformer fusion network for Alzheimer's disease diagnosis. *Int J Imaging Syst Technol* 2024; 34: e22970.
  75. Nguyen HH, Blaschko MB and Saarakkala S, et al. Clinically-inspired multi-agent transformers for disease trajectory forecasting from multimodal data. *IEEE Trans Med Imaging* 2023; 43: 529–541.
  76. Mahmood T, Rehman A and Saba T, et al. Alzheimer's disease unveiled: cutting-edge multi-modal neuroimaging and computational methods for enhanced diagnosis. *Biomed Signal Process Control* 2024; 97: 106721.
  77. Wang Y, Gao R and Wei T, et al. Predicting long-term progression of Alzheimer's disease using a multimodal deep learning model incorporating interaction effects. *J Transl Med* 2024; 22: 265.
  78. Houria L, Belkhamisa N and Cherfa A, et al. Multi-modality MRI for Alzheimer's disease detection using deep learning. *Phys Eng Sci Med* 2022; 45: 1043–1053.
  79. Xu H, Zhong S and Zhang Y. Multi-level fusion network for mild cognitive impairment identification using multi-modal neuroimages. *Phys Med Biol* 2023; 68: 095018.
  80. He Y, Wu J and Zhou L, et al. Quantification of cognitive function in Alzheimer's disease based on deep learning. *Front Neurosci* 2021; 15: 651920.
  81. Jia H, Wang Y and Duan Y, et al. Alzheimer's disease classification based on image transformation and features fusion. *Comput Math Methods Med* 2021; 2021: 9624269.
  82. Jia H and Lao H. Deep learning and multimodal feature fusion for the aided diagnosis of Alzheimer's disease. *Neural Comput Appl* 2022; 34: 19585–19598.
  83. Castellazzi G, Cuzzoni MG and Cotta Ramusino M, et al. A machine learning approach for the differential diagnosis of Alzheimer and vascular dementia fed by mri selected features. *Front Neuroinform* 2020; 14: 25.
  84. Meng X, Wei Q and Meng L, et al. Feature fusion and detection in Alzheimer's disease using a novel genetic multi-kernel svm based on MRI imaging and gene data. *Genes* 2022; 13: 837.
  85. Jiao Z, Chen S and Shi H, et al. Multi-modal feature selection with feature correlation and feature structure fusion for MCI and AD classification. *Brain Sci* 2022; 12: 80.

86. Ban Y, Lao H and Li B, et al. Diagnosis of Alzheimer's disease using hypergraph p-Laplacian regularized multi-task feature learning. *J Biomed Inform* 2023; 140: 104326.
87. Liu Y, Wang L and Ning X, et al. Enhancing early Alzheimer's disease classification accuracy through the fusion of SMRI and RSMEG data: a deep learning approach. *Front Neurosci* 2024; 18: 1480871.
88. Li Z, Dong A and Zhou J. Research of low-rank representation and discriminant correlation analysis for Alzheimer's disease diagnosis. *Comput Math Methods Med* 2020; 2020: 5294840.
89. Guelib B, Zarour K and Hermessi H, et al. Same-subject-modalities-interactions: a novel framework for MRI and PET multi-modality fusion for Alzheimer's disease classification. *IEEE Access* 2023; 11: 48715–48738.
90. Wang B, Li L and Peng L, et al. Multigroup recognition of dementia patients with dynamic brain connectivity under multimodal cortex parcellation. *Biomed Signal Process Control* 2022; 76: 103725.
91. Lu D, Popuri K and Ding GW, et al. Multimodal and multi-scale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Sci Rep* 2018; 8: 5697.
92. Nai YH, Liu H and Reilhac A, et al. Validation of deep learning-based nonspecific estimates for amyloid burden quantification with longitudinal data. *Phys Med* 2022; 99: 85–93.
93. Goenka N and Tiwari S. Alzheimer's detection using various feature extraction approaches using a multimodal multi-class deep learning model. *Int J Imaging Syst Technol* 2023; 33: 588–609.
94. Syed ZS, Syed MSS and Lech M, et al. Automated recognition of Alzheimer's dementia using bag-of-deep-features and model ensembling. *IEEE Access* 2021; 9: 88377–88390.
95. Feng Y, Kim M and Yao X, et al. Deep multiview learning to identify imaging-driven subtypes in mild cognitive impairment. *BMC Bioinformatics* 2022; 23: 402.
96. Xue L, Fu Y and Gao X, et al. [18f] FDG PET integrated with structural MRI for accurate brain age prediction. *Eur J Nucl Med Mol Imaging* 2024; 51: 3617–3629.
97. Sheng J, Wang B and Zhang Q, et al. A novel joint HCPMMP method for automatically classifying Alzheimer's and different stage MCI patients. *Behav Brain Res* 2019; 365: 210–221.
98. Liu M, Cheng D and Wang K, et al. Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis. *Neuroinformatics* 2018; 16: 295–308.
99. Hao X, Bao Y and Guo Y, et al. Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease. *Med Image Anal* 2020; 60: 101625.
100. Xu X, Wang T and Li W, et al. Morphological, structural, and functional networks highlight the role of the cortical-subcortical circuit in individuals with subjective cognitive decline. *Front Aging Neurosci* 2021; 13: 688113.
101. Leng Y, Cui W and Peng Y, et al. Multimodal cross enhanced fusion network for diagnosis of Alzheimer's disease and subjective memory complaints. *Comput Biol Med* 2023; 157: 106788.
102. Herzog R, Rosas FE and Whelan R, et al. Genuine high-order interactions in brain networks and neurodegeneration. *Neurobiol Dis* 2022; 175: 105918.
103. Abdelaziz M, Wang T and Elazab A. Fusing multimodal and anatomical volumes of interest features using convolutional auto-encoder and convolutional neural networks for Alzheimer's disease diagnosis. *Front Aging Neurosci* 2022; 14: 812870.
104. Qiang YR, Zhang SW and Li JN, et al. Diagnosis of Alzheimer's disease by joining dual attention CNN and MLP based on structural MRIS, clinical and genetic data. *Artif Intell Med* 2023; 145: 102678.
105. Zhang J, He X and Liu Y, et al. Multi-modal cross-attention network for Alzheimer's disease diagnosis with multi-modality data. *Comput Biol Med* 2023; 162: 107050.
106. Pena D, Suescun J and Schiess M, et al. Toward a multi-modal computer-aided diagnostic tool for Alzheimer's disease conversion. *Front Neurosci* 2022; 15: 744190.
107. Kim J and Lee B. Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine. *Hum Brain Mapp* 2018; 39: 3728–3741.
108. Peng J, Wang W and Song Q, et al. 18f-FDG-PET radiomics based on white matter predicts the progression of mild cognitive impairment to Alzheimer disease: a machine learning study. *Acad Radiol* 2023; 30: 1874–1884.
109. Zhang Q, Du Q and Liu G. A whole-process interpretable and multi-modal deep reinforcement learning for diagnosis and analysis of Alzheimer's disease\*. *J Neural Eng* 2021; 18: 066032.
110. Martinez-Murcia FJ, Ortiz A and Gorriz JM, et al. Studying the manifold structure of Alzheimer's disease: a deep learning approach using convolutional autoencoders. *IEEE J Biomed Health Inform* 2019; 24: 17–26.
111. Lee G, Kang B and Nho K, et al. Mildint: deep learning-based multimodal longitudinal data integration framework. *Front Genet* 2019; 10: 617.
112. Mehdiipour Ghazi M, Selnes P and Timón-Reina S, et al. Comparative analysis of multimodal biomarkers for amyloid-beta positivity detection in Alzheimer's disease cohorts. *Front Aging Neurosci* 2024; 16: 1345417.
113. Wang Y, Xu C and Park JH, et al. Diagnosis and prognosis of Alzheimer's disease using brain morphometry and white matter connectomes. *NeuroImage: Clin* 2019; 23: 101859.
114. Tabarestani S, Aghili M and Eslami M, et al. A distributed multitask multimodal approach for the prediction of Alzheimer's disease in a longitudinal study. *NeuroImage* 2020; 206: 116317.
115. Ostertag C, Visani M and Urruty T, et al. Long-term cognitive decline prediction based on multi-modal data using multimodal3dsiamesenet: transfer learning from

- Alzheimer's disease to Parkinson's disease. *Int J Comput Assist Radiol Surg* 2023; 18: 809–818.
116. Tabarestani S, Eslami M and Cabrerizo M, et al. A tensorized multitask deep learning network for progression prediction of Alzheimer's disease. *Front Aging Neurosci* 2022; 14: 810873.
  117. Zhu Q, Xu B and Huang J, et al. Deep multi-modal discriminative and interpretability network for Alzheimer's disease diagnosis. *IEEE Trans Med Imaging* 2022; 42: 1472–1483.
  118. Bhagwat N, Viviano JD and Voineskos AN, et al. Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLoS Comput Biol* 2018; 14: e1006376.
  119. Bazargani JS, Rahim N and Sadeghi-Niaraki A, et al. Alzheimer's disease diagnosis in the metaverse. *Comput Methods Programs Biomed* 2024; 255: 108348.
  120. Gaeta AM, Quijada-López M and Barbé F, et al. Predicting Alzheimer's disease CSF core biomarkers: a multimodal machine learning approach. *Front Aging Neurosci* 2024; 16: 1369545.
  121. Zhou T, Thung KH and Zhu X, et al. Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Hum Brain Mapp* 2019; 40: 1001–1016.
  122. Chen Y, Su Y and Wu J, et al. Combining blood-based biomarkers and structural MRI measurements to distinguish persons with and without significant amyloid plaques. *J Alzheimers Dis* 2024; 98: 1415–1426.
  123. Yuan S, Li H and Wu J, et al. Classification of mild cognitive impairment with multimodal data using both labeled and unlabeled samples. *IEEE/ACM Trans Comput Biol Bioinform* 2021; 18: 2281–2290.
  124. Sun X, Guo W and Shen J. Toward attention-based learning to predict the risk of brain degeneration with multimodal medical data. *Front Neurosci* 2023; 16: 1043626.
  125. Perez-Gonzalez J, Jiménez-Ángeles L and Saavedra KR, et al. Mild cognitive impairment classification using combined structural and diffusion imaging biomarkers. *Phys Med Biol* 2021; 66: 155010.
  126. Sheng X, Cai H and Nie Y, et al. Modality-aware discriminative fusion network for integrated analysis of brain imaging genomics. *IEEE Trans Neural Netw Learn Syst* 2024; 36: 8577–8591.
  127. Xu E, Zhang J and Li J, et al. Pathology steered stratification network for subtype identification in Alzheimer's disease. *Med Phys* 2024; 51: 1190–1202.
  128. Xue C, Kowshik SS and Lteif D, et al. AI-based differential diagnosis of dementia etiologies on multimodal data. *Nat Med* 2024; 30: 2977–2989.
  129. Fu Y, Xu L and Zhang Y, et al. Classification and diagnosis model for Alzheimer's disease based on multimodal data fusion. *Medicine* 2024; 103: e41016.
  130. Olatunde OO, Oyetunde KS and Han J, et al. Multiclass classification of Alzheimer's disease prodromal stages using sequential feature embeddings and regularized multikernel support vector machine. *NeuroImage* 2024; 304: 120929.
  131. Yuan Z, Li X and Hao Z, et al. Intelligent prediction of Alzheimer's disease via improved multifeature squeeze-and-excitation-dilated residual network. *Sci Rep* 2024; 14: 11994.
  132. Akkaya UM and Kalkan H. A new approach for multimodal usage of gene expression and its image representation for the detection of Alzheimer's disease. *Biomolecules* 2023; 13: 1563.
  133. Qiu S, Joshi PS and Miller MI, et al. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain* 2020; 143: 1920–1933.
  134. Eslami M, Tabarestani S and Adjouadi M. A unique color-coded visualization system with multimodal information fusion and deep learning in a longitudinal study of Alzheimer's disease. *Artif Intell Med* 2023; 140: 102543.
  135. Gupta Y, Lama RK and Kwon GR, et al. Prediction and classification of Alzheimer's disease based on combined features from apolipoprotein-e genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers. *Front Comput Neurosci* 2019; 13: 72.
  136. Kumari R, Nigam A and Pushkar S. An efficient combination of quadruple biomarkers in binary classification using ensemble machine learning technique for early onset of Alzheimer disease. *Neural Comput Appl* 2022; 34: 11865–11884.
  137. Zarei A, Keshavarz A and Jafari E, et al. Automated classification of Alzheimer's disease, mild cognitive impairment, and cognitively normal patients using 3D convolutional neural network and radiomic features from t1-weighted brain MRI: a comparative study on detection accuracy. *Clin Imaging* 2024; 115: 110301.
  138. Yang L, Wang X and Guo Q, et al. Deep learning based multimodal progression modeling for Alzheimer's disease. *Stat Biopharm Res* 2021; 13: 337–343.
  139. Abdelaziz M, Wang T and Elazab A. Alzheimer's disease diagnosis framework from incomplete multimodal data using convolutional neural networks. *J Biomed Inform* 2021; 121: 103863.
  140. AlMohimeed A, Saad RM and Mostafa S, et al. Explainable artificial intelligence of multi-level stacking ensemble for detection of Alzheimer's disease based on particle swarm optimization and the sub-scores of cognitive biomarkers. *IEEE Access* 2023; 11: 123173.
  141. Meng L and Zhang Q. Research on early diagnosis of Alzheimer's disease based on dual fusion cluster graph convolutional network. *Biomed Signal Process Control* 2023; 86: 105212.
  142. Venugopalan J, Tong L and Hassanzadeh HR, et al. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Sci Rep* 2021; 11: 3254.
  143. Yu X, Liu J and Lu Y, et al. Early diagnosis of Alzheimer's disease using a group self-calibrated coordinate attention

- network based on multimodal MRI. *Sci Rep* 2024; 14: 24210.
144. Yin Y, Wang H and Liu S, et al. Internet of things for diagnosis of Alzheimer's disease: a multimodal machine learning approach based on eye movement features. *IEEE Int Things J* 2023; 10: 11476–11485.
  145. Zhang C, Lei X and Ma W, et al. Diagnosis framework for probable Alzheimer's disease and mild cognitive impairment based on multi-dimensional emotion features. *J Alzheimers Dis* 2024; 97: 1125–1137.
  146. De A and Chowdhury AS. DTI based Alzheimer's disease classification with rank modulated fusion of CNNs and random forest. *Expert Syst Appl* 2021; 169: 114338.
  147. Chen Y, Wang H and Zhang D, et al. Multi-feature fusion learning for Alzheimer's disease prediction using EEG signals in resting state. *Front Neurosci* 2023; 17: 1272834.
  148. El-Sappagh S, Abuhmed T and Alouffi B, et al. The role of medication data to enhance the prediction of Alzheimer's progression using machine learning. *Comput Intell Neurosci* 2021; 2021: 8439655.
  149. Ilias L and Askounis D. Multimodal deep learning models for detecting dementia from speech and transcripts. *Front Aging Neurosci* 2022; 14: 830943.
  150. Yiğit A, Baştanlar Y and Işık Z. Dementia diagnosis by ensemble deep neural networks using FDG-PET scans. *Signal Image Video Process* 2022; 16: 2203–2210.
  151. Boudaya A, Chaabene S and Bouaziz B, et al. Mild cognitive impairment detection based on EEG and HRV data. *Digit Signal Process* 2024; 147: 104399.
  152. Franciotti R, Nardini D and Russo M, et al. Comparison of machine learning-based approaches to predict the conversion to Alzheimer's disease from mild cognitive impairment. *Neuroscience* 2023; 514: 143–152.
  153. Mohan Gowda V, Arakeri MP and Raghu Ram Prasad V. Multimodal classification technique for fall detection of Alzheimer's patients by integration of a novel piezoelectric crystal accelerometer and aluminum gyroscope with vision data. *Adva Mater Sci Eng* 2022; 2022: 9258620.
  154. Vaghari D, Kabir E and Henson RN. Late combination shows that MEG adds to MRI in classifying MCI versus controls. *Neuroimage* 2022; 252: 119054.
  155. Cirincione A, Lynch K and Bennett J, et al. Prediction of future dementia among patients with mild cognitive impairment (MCI) by integrating multimodal clinical data. *Heliyon* 2024; 10: e36728.
  156. Wang Y, Liu X and Yu C. Assisted diagnosis of Alzheimer's disease based on deep learning and multimodal feature fusion. *Complexity* 2021; 2021: 6626728.
  157. Yan Y, Somer E and Grau V. Classification of amyloid pet images using novel features for early diagnosis of Alzheimer's disease and mild cognitive impairment conversion. *Nucl Med Commun* 2019; 40: 242–248.
  158. El-Gamal FEZA, Elmogy MM and Khalil A, et al. Personalized computer-aided diagnosis for mild cognitive impairment in Alzheimer's disease based on SMRI and 11c PIB-PET analysis. *IEEE Access* 2020; 8: 218982.
  159. Chen Z, Chen K and Li Y, et al. Structural, static, and dynamic functional MRI predictors for conversion from mild cognitive impairment to Alzheimer's disease: inter-cohort validation of Shanghai memory study and ADNI. *Hum Brain Mapp* 2024; 45: e26529.
  160. Zhang Y, He X and Chan YH, et al. Multi-modal graph neural network for early diagnosis of Alzheimer's disease from SMRI and PET scans. *Comput Biol Med* 2023; 164: 107328.
  161. Cheng S, Shang P and Zhang Y, et al. An fnirs representation and fnirs-scales multimodal fusion method for auxiliary diagnosis of amnesic mild cognitive impairment. *Biomed Signal Process Control* 2024; 96: 106646.
  162. Arco JE, Ramírez J and Górriz JM, et al. Data fusion based on searchlight analysis for the prediction of Alzheimer's disease. *Expert Syst Appl* 2021; 185: 115549.
  163. Zhang H, Ni M and Yang Y, et al. Patch-based interpretable deep learning framework for Alzheimer's disease diagnosis using multimodal data. *Biomed Signal Process Control* 2025; 100: 107085.
  164. Biswas R and Gini JR. Multi-class classification of Alzheimer's disease detection from 3D MRI image using ML techniques and its performance analysis. *Multimed Tools Appl* 2024; 83: 33527–33554.
  165. Lakhan A, Mohammed MA and Abd Ghani MK, et al. Fdcnn-as: Federated deep convolutional neural network Alzheimer detection schemes for different age groups. *Inf Sci (Ny)* 2024; 677: 120833.
  166. Alty J, Goldberg LR and Roccati E, et al. Development of a smartphone screening test for preclinical Alzheimer's disease and validation across the dementia continuum. *BMC Neurol* 2024; 24: 127.
  167. Martinc M, Haider F and Pollak S, et al. Temporal integration of text transcripts and acoustic features for Alzheimer's diagnosis based on spontaneous speech. *Front Aging Neurosci* 2021; 13: 642647.
  168. Li Y, Chen G and Wang G, et al. Dominating Alzheimer's disease diagnosis with deep learning on SMRI and DTI-MD. *Front Neurol* 2024; 15: 1444795.
  169. Zhang M, Cui Q and Lü Y, et al. A multimodal learning machine framework for Alzheimer's disease diagnosis based on neuropsychological and neuroimaging data. *Comput Indust Eng* 2024; 197: 110625.
  170. Fang X, Liu Z and Xu M. Ensemble of deep convolutional neural networks based multi-modality images for Alzheimer's disease diagnosis. *IET Image Process* 2020; 14: 318–326.
  171. Zhang T and Shi M. Multi-modal neuroimaging feature fusion for diagnosis of Alzheimer's disease. *J Neurosci Methods* 2020; 341: 108795.
  172. Gao X, Shi F and Shen D, et al. Task-induced pyramid and attention gan for multimodal brain image imputation and

- classification in Alzheimer's disease. *IEEE J Biomed Health Inform* 2021; 26: 36–43.
173. Lin W, Lin W and Chen G, et al. Bidirectional mapping of brain MRI and pet with 3D reversible gan for the diagnosis of Alzheimer's disease. *Front Neurosci* 2021; 15: 646013.
  174. Choi H and Lee DS. Generation of structural MR images from amyloid pet: application to MR-less quantification. *J Nucl Med* 2018; 59: 1111–1117.
  175. Liu H, Nai YH and Saridin F, et al. Improved amyloid burden quantification with nonspecific estimates using deep learning. *Eur J Nucl Med Mol Imaging* 2021; 48: 1842–1853.
  176. Zhang M, Sun L and Kong Z, et al. Pyramid-attentive gan for multimodal brain image complementation in Alzheimer's disease classification. *Biomed Signal Process Control* 2024; 89: 105652.
  177. Liu Y, Yue L and Xiao S, et al. Assessing clinical progression from subjective cognitive decline to mild cognitive impairment with incomplete multi-modal neuroimages. *Med Image Anal* 2022; 75: 102266.
  178. Chen K, Weng Y and Hosseini AA, et al. A comparative study of GNN and MLP based machine learning for the diagnosis of Alzheimer's disease involving data synthesis. *Neural Netw* 2024; 169: 442–452.
  179. Gao X, Shi F and Shen D, et al. Multimodal transformer network for incomplete image generation and diagnosis of Alzheimer's disease. *Comput Med Imaging Graph* 2023; 110: 102303.
  180. Hwang U, Kim SW and Jung D, et al. Real-world prediction of preclinical Alzheimer's disease with a deep generative model. *Artif Intell Med* 2023; 144: 102654.
  181. Nguyen M, He T and An L, et al. Predicting Alzheimer's disease progression using deep recurrent neural networks. *NeuroImage* 2020; 222: 117203.
  182. Fedorov A, Geenjaar E and Wu L, et al. Self-supervised multimodal learning for group inferences from MRI data: discovering disorder-relevant brain regions and multimodal links. *NeuroImage* 2024; 285: 120485.
  183. Martí-Juan G, Lorenzi M and Piella G, et al. Mc-rvae: Multi-channel recurrent variational autoencoder for multimodal Alzheimer's disease progression modelling. *NeuroImage* 2023; 268: 119892.
  184. Chen J, Wang Y and Zeb A, et al. Multimodal mixing convolutional neural network and transformer for Alzheimer's disease recognition. *Expert Syst Appl* 2025; 259: 125321.
  185. Zhang F, Li Z and Zhang B, et al. Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease. *Neurocomputing* 2019; 361: 185–195.
  186. Chandrashekar PB, Alatkar S and Wang J, et al. Deepgami: deep biologically guided auxiliary learning for multimodal integration and imputation to improve genotype–phenotype prediction. *Genome Med* 2023; 15: 88.
  187. Rahim N, El-Sappagh S and Ali S, et al. Prediction of Alzheimer's progression based on multimodal deep-learning-based fusion and visual explainability of time-series data. *Information Fusion* 2023; 92: 363–388.
  188. Tu Y, Lin S and Qiao J, et al. Alzheimer's disease diagnosis via multimodal feature fusion. *Comput Biol Med* 2022; 148: 105901.
  189. Tang Y, Xiong X and Tong G, et al. Multimodal diagnosis model of Alzheimer's disease based on improved transformer. *Biomed Eng Online* 2024; 23: 8.
  190. El-Sappagh S, Ali F and Abuhmed T, et al. Automatic detection of Alzheimer's disease progression: an efficient information fusion approach with heterogeneous ensemble classifiers. *Neurocomputing* 2022; 512: 203–224.
  191. Bo W, Sullivan SS and Zhang X, et al. A telemedicine analytic framework for fully and semi-automatic Alzheimer's disease screening using clock drawing test. *IEEE J Biomed Health Inform* 2024; 28: 7503–7516.
  192. Saleh H, Amer E and Abuhmed T, et al. Computer aided progression detection model based on optimized deep LSTM ensemble model and the fusion of multivariate time series data. *Sci Rep* 2023; 13: 16336.
  193. Balaji P, Chaurasia MA and Bilfaqih SM, et al. Hybridized deep learning approach for detecting Alzheimer's disease. *Biomedicines* 2023; 11: 149.
  194. Wilson RS, Leurgans SE and Boyle PA, et al. Cognitive decline in prodromal Alzheimer disease and mild cognitive impairment. *Arch Neurol* 2011; 68: 351–356.
  195. Gharehbaghi A and Babic A. Artificial intelligence in cognitive decline diagnosis: evaluating cutting-edge techniques and modalities. In: 23rd International Conference on Informatics, Management, and Technology in Healthcare, 328, pp.46–50.
  196. Punjabi A, Martersteck A and Wang Y, et al. Neuroimaging modality fusion in Alzheimer's classification using convolutional neural networks. *PLoS ONE* 2019; 14: e0225759.
  197. Fabietti M, Mahmud M and Lotfi A, et al. Early detection of Alzheimer's disease from cortical and hippocampal local field potentials using an ensembled machine learning model. *IEEE Trans Neural Syst Rehabil Eng* 2023; 31: 2839–2848.
  198. Houria L, Belkhamisa N and Cherfa A, et al. Multimodal magnetic resonance imaging for Alzheimer's disease diagnosis using hybrid features extraction and ensemble support vector machines. *Int J Imaging Syst Technol* 2023; 33: 610–621.