IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning Approaches in Alzheimer's Disease: Early Diagnosis, Prognosis, and Treatment Optimisation

Vatsal Choksi

Student, Bhagwan Mahavir College of Computer Application.

Dr. Sneha Patel

Assistant Professor, Bhagwan Mahavir College of Computer Application.

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that has emerged as a significant global health challenge, affecting millions worldwide. It is characterised by cognitive decline, memory loss, and behavioural changes, significantly burdening patients, caregivers, and healthcare systems. Early diagnosis and effective intervention remain critical yet challenging, given the complex nature of the disease and its heterogeneous progression. Traditional diagnostic methods rely on invasive, expensive, and often time-consuming procedures, underscoring the urgent need for innovative diagnostic and therapeutic strategies.

Machine learning (ML), a subset of artificial intelligence, has revolutionised medical research by enabling the analysis of complex, high-dimensional datasets. In Alzheimer's research, ML has shown immense potential to enhance diagnostic accuracy, predict disease progression, and personalise treatment plans. ML models leverage data from various sources, including neuroimaging, genomics, biomarkers, and clinical records, to identify subtle patterns and accurately predict outcomes. Applying supervised learning, deep learning techniques such as convolutional neural networks (CNNs) to imaging data, and unsupervised clustering algorithms for patient stratification has significantly advanced the field.

This paper focuses on three critical areas: (1) early diagnosis of AD through ML-based analysis of neuroimaging and biomarkers, (2) prediction of disease progression using longitudinal data and predictive models, and (3) optimisation of treatment plans and personalised care through reinforcement learning and predictive analytics. The study also explores the challenges associated with ML applications, such as data heterogeneity, interpretability, and ethical concerns, and discusses potential solutions to address these barriers.

In conclusion, machine learning represents a transformative tool in Alzheimer's research. It promises to improve early detection, enhance disease management, and pave the way for precision medicine. Its integration into clinical practice can potentially significantly mitigate the global burden of AD.

Keywords: *convolutional neural networks, Machine learning, Alzheimer's disease, neuroimaging, genomics, biomarkers.*

1. Introduction

1.1 Overview of Alzheimer's Disease

Definition and Symptoms

Alzheimer's disease (AD) is a chronic neurodegenerative disorder primarily affecting older adults, characterised by progressive cognitive decline and memory loss. It is the most common cause of dementia, accounting for approximately 60-80% of all dementia cases. The disease manifests through a wide range of symptoms, including:

- **Early-stage symptoms**: Mild memory lapses, difficulty recalling recent events, and trouble finding the right words.
- **Moderate-stage symptoms**: Increased confusion, challenges in performing routine tasks, mood swings, and personality changes.
- **Late-stage symptoms**: Severe memory loss, difficulty recognising loved ones, and a decline in physical abilities, including walking and swallowing.

Stages of AD

AD progresses in three main stages:

1. **Mild (early-stage)**: Subtle memory lapses and minor functional impairments.
2. **Moderate (middle-stage)**: Significant cognitive decline, noticeable behavioural changes, and reduced independence.
3. **Severe (late-stage)**: Complete dependence on caregivers and severe physical and cognitive deterioration.

Global Prevalence and Socio-Economic Impact

According to the World Health Organisation (WHO), approximately 55 million people were living with dementia in 2021, with AD being the leading contributor. The number is projected to triple by 2050 due to ageing populations. The socio-economic impact of AD is profound, including:

- Direct costs of care (medical, institutional, and home-based).
- Indirect costs, such as loss of productivity among caregivers.
- Emotional and psychological toll on families and communities.

Challenges in Early Diagnosis and Treatment

- Current diagnostic methods (e.g., neuroimaging, cerebrospinal fluid testing) are expensive, invasive, and inaccessible.
- Many patients are diagnosed at moderate or late stages when treatment options are less effective.
- Existing treatments focus primarily on symptom management rather than modifying the disease's progression.

1.2 Role of Machine Learning in Healthcare

Overview of ML Applications in Medical Research

Machine learning (ML), a branch of artificial intelligence (AI), enables computers to learn from data and make predictions without being explicitly programmed. ML has transformed healthcare by providing tools to analyse large, complex datasets and uncover insights that traditional methods cannot. Applications include:

- Predicting patient outcomes and risk stratification.
- Analysing medical imaging for disease detection (e.g., cancers, brain disorders).

- Drug discovery and optimisation of clinical trials.

ML as a Tool for Complex Dataset Analysis

In Alzheimer's research, ML is particularly valuable because it can process multi-modal datasets, such as:

- **Neuroimaging data**: MRI and PET scans to detect brain changes indicative of AD.
- **Genomic data**: Identifying genetic predispositions (e.g., APOE-ε4 allele).
- **Clinical and behavioural data**: Patterns in cognitive test scores and patient history.
- **Biomarker data**: Analysis of amyloid-beta and tau protein levels in cerebrospinal fluid or blood.

Importance of ML in Overcoming Challenges in AD Research

- **Early diagnosis**: Identifying preclinical AD stages by detecting subtle changes in imaging and biomarkers.
- **Disease progression modelling**: Predicting the rate at which a patient's condition may worsen.
- **Personalised treatment**: Developing targeted therapies based on individual data.
- **Reducing costs and time**: Automating analyses that traditionally require significant resources.

1.3 Objectives of the Study

Exploring the Role of ML in Early Diagnosis

The study aims to investigate how ML algorithms can enhance the accuracy and timeliness of AD diagnosis by analysing data from various sources, including neuroimaging, genomics, and biomarkers. This includes applying supervised learning to classification tasks and deep learning to interpret imaging data.

Understanding ML-Based Models for Disease Progression Prediction

Disease progression in AD varies across individuals. This study explores the use of predictive models, such as regression and time-series analyses, to forecast cognitive and functional decline, enabling clinicians to plan interventions more effectively.

Investigating Personalised Treatment Approaches Using ML

This objective uses ML techniques, such as reinforcement learning and predictive analytics, to optimise treatment strategies. By integrating patient-specific data, ML can help identify the most effective therapies and monitor treatment responses in real time, paving the way for precision medicine in AD care.

2.2 Data Sources for Alzheimer's Disease Research

OASIS (Open Access Series of Imaging Studies)

The OASIS dataset is a publicly available resource that significantly contributes to Alzheimer's disease (AD) research. It includes high-resolution neuroimaging data, such as MRI scans, and is widely used for training and evaluating machine learning (ML) models. The dataset is particularly valuable due to:

- **Longitudinal data**: Includes imaging from patients at different time points, enabling researchers to study disease progression.
- **Diversity**: Covers a wide range of subjects, including healthy individuals and patients across the AD spectrum (from mild cognitive impairment to severe stages).
- **Accessibility**: The open-access nature facilitates collaboration and innovation in AD research.

Key applications of OASIS data include early AD detection, the study of structural brain changes (e.g., hippocampal atrophy), and the development of predictive models for cognitive decline.

Kaggle Dataset

The Kaggle platform hosts various Alzheimer 's-related datasets, often contributed by research institutions or generated during data science competitions. Standard features of these datasets include:

- **Neuroimaging**: MRI and PET scan data for brain analysis.
- **Clinical records**: Patient demographics, cognitive test scores, and family history of AD.
- **Biomarkers**: Information on amyloid-beta, tau proteins, and other AD-related biomarkers.
- **Genomic data**: Details on genetic predispositions, such as the APOE-ε4 allele.

Kaggle datasets are particularly advantageous for data science projects due to their ease of access and extensive pre-processed formats. However, these datasets may sometimes lack detailed clinical metadata, requiring careful contextual interpretation.

Clinical Datasets and Their Challenges

Clinical datasets, such as those from hospitals or extensive research studies, are essential for developing real-world ML models. Examples include datasets from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the National Alzheimer's Coordinating Centre (NACC). However, analysing clinical datasets comes with several challenges:

- **Heterogeneity**: Variability in data collection methods, patient demographics, and clinical protocols across institutions.
- **Missing Data**: Clinical datasets often have incomplete records due to inconsistent follow-ups or data entry errors, requiring Imputation techniques.
- **Ethical and Privacy Concerns**: Sharing and accessing clinical data are subject to stringent regulations (e.g., HIPAA, GDPR), which can limit the availability of large-scale datasets.
- **Limited Standardisation**: Differences in diagnostic criteria, imaging protocols, and biomarker measurements complicate study comparisons.

Wearable Devices and Patient-Generated Health Data

Advances in wearable technology have introduced new opportunities to collect real-time, patient-generated health data for Alzheimer's research. Wearable devices, such as smartwatches and fitness trackers, can measure:

- **Cognitive performance**: Using gamified cognitive tests or reaction time assessments.
- **Behavioural patterns**: Monitoring sleep quality, physical activity, and social interactions to identify early signs of AD.
- **Vital signs**: Heart rate variability and other physiological markers linked to AD progression.

Challenges in wearable data include:

- **Data Volume**: Continuous monitoring generates vast data, requiring scalable storage and processing solutions.
- **Noise and Artefacts**: Wearable sensors are prone to inaccuracies due to user behaviour (e.g., improper usage) or device limitations.
- **Integration with Clinical Data**: Combining wearable data with traditional biomarkers and imaging data for a holistic understanding of AD remains complex.

2.3 Challenges in Analysing Biomarker Data

High Dimensionality of Data

Biomarker datasets, especially those derived from genomics, proteomics, or neuroimaging, often contain thousands of variables per sample. The "curse of dimensionality" arises when the number of features far exceeds the number of samples, making it challenging to train accurate ML models without overfitting. Dimensionality reduction techniques, such as Principal Component Analysis (PCA) and autoencoders, are often employed to address this issue.

Variability Across Patients

Biomarker expression levels can vary significantly between individuals due to genetic, environmental, and lifestyle factors. This variability complicates the development of generalised ML models and necessitates approaches like:

- **Stratified analysis**: Grouping patients based on demographic or genetic similarities.
- **Ensemble learning**: Combining predictions from multiple models to improve robustness.

Need for Advanced Algorithms for Feature Selection and Interpretation

Feature selection is critical in biomarker analysis to identify the most relevant predictors of AD onset and progression. Challenges include:

- **Biological Relevance**: ML algorithms often select features based on statistical significance, which may not always align with biological relevance.
- **Interpretability**: Black-box models like deep neural networks lack transparency, making it difficult to understand the role of specific biomarkers.
- **Computational Complexity**: Processing large datasets with complex biomarker interactions requires high-performance computing resources.

Emerging methods, such as explainable AI (XAI) and feature importance scoring (e.g., SHAP values), are helping researchers address these challenges, improving the reliability and interpretability of ML-based biomarker analysis.

3. Two CNN architectures and visualising their performance.

The execution begins by creating instances of both models—the basic_cnn and advanced_cnn—using their respective creation functions. These models are then trained and evaluated using the train_and_evaluate_model function, with results stored in the basic_results and advanced_results variables. The model names are passed as parameters for precise identification in the output.

The results analysis is conducted through two visualisation approaches:

1. First, it calls the plot_results and print_comparison functions to generate performance metrics and training curves
2. Second, it creates detailed confusion matrices for both models using Seaborn's heatmap visualisation

The confusion matrix visualisation is particularly sophisticated, using Matplotlib's subplots to create a side-by-side comparison (1x2 grid with a 15x6 figure size). The class names are extracted from the training generator's indices to ensure accurate labelling. For each model, a confusion matrix is computed using sci-kit-learn's confusion_matrix function and visualised using Seaborn's heatmap with the following features:

- Blue colour scheme (cmap='Blues')
- Annotated values (annot=True)
- Decimal format ('d' format)
- Proper axis labels ('True Label' and 'Predicted Label')
- Clear titles distinguishing between Basic and Advanced CNN results

This visualisation approach enables direct comparison of model performance across different classes, making it easy to identify:

- Where each model excels or struggles
- Patterns in misclassifications
- Overall classification accuracy
- Class-specific performance differences

3.1Basic CNN:

- Trained for 26 epochs out of the planned 50 (stopped early due to no improvement)
- Started with an accuracy of 41.37% and steadily improved to 88.21%
- Learning rate was adaptively reduced from 0.001 to 0.00004, showing proper learning rate scheduling
- Training time: 9,364.22 seconds (approximately 2.6 hours)
- Best validation accuracy: 71.92% (Epoch 14)
- Showed signs of overfitting with fluctuating validation accuracy

3.2 Advanced CNN (ResNet50V2):

- Trained for ten epochs before early stopping
- Initial accuracy: 45.64%, final accuracy: 66.62%
- More stable validation accuracy but lower overall performance
- Training time: 4,914.92 seconds (approximately 1.4 hours)
- Best validation accuracy: 59.96%

Final Test Results:

Basic CNN:

- Test Accuracy: 56.06%
- Strong performance in specific categories:
  - Perfect precision (1.00) for Moderate Impairment
  - High precision (0.88) for No Impairment
  - Good recall (0.85) for Very Mild Impairment
- Weighted average F1-score: 0.55

Advanced CNN:

- Test Accuracy: 49.73%
- More balanced but generally lower performance:
  - Best performance in No Impairment category (0.60 precision, 0.76 recall)
  - Struggled with Moderate Impairment (0.07 precision)
  - Weighted average F1-score: 0.46

Key Observations:

1. **Model Efficiency:**

The advanced CNN trained faster despite its complexity (47.5% less time)
Basic CNN achieved better accuracy but required more training time

2. **Class-specific Performance:**

Both models handled "No Impairment" cases relatively well
Significant challenges with the "Moderate Impairment" classification
Basic CNN showed better discrimination between classes

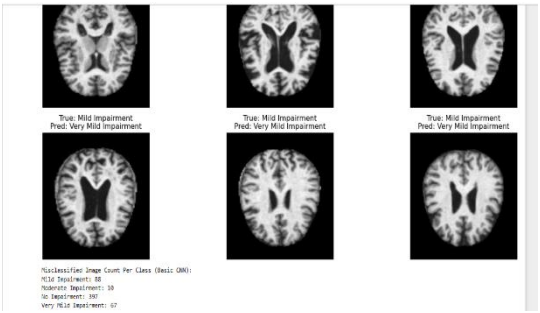
3. **Clinical Implications:**

A basic CNN might be more reliable for initial screening
High false positive rate in both models suggests a need for improvement
Models show promise but require refinement for clinical application

4. **Technical Notes:**

Warning messages indicate potential optimisation opportunities in the code
HDF5 format warnings suggest updating to the newer Keras model saving format
Early stopping effectively prevented overfitting in both models

The results suggest that while both models show promise, neither achieves the level of accuracy typically required for medical diagnostics. The basic CNN outperformed the advanced model despite its simpler architecture, indicating that complex architectures don't always yield better results. Future improvements might focus on data augmentation, architecture optimisation, or ensemble methods to improve accuracy while maintaining reasonable training times.



The Basic CNN demonstrates remarkable precision in specific categories, particularly achieving perfect precision (1.00) for Moderate Impairment and high precision (0.88) for No Impairment cases. However, this high precision comes at a cost, as the model shows inconsistent recall across different impairment levels. This suggests that while the Basic CNN is highly confident in its predictions, it may miss several cases, particularly in the Moderate Impairment category.

In contrast, the Advanced CNN exhibits more balanced performance metrics across all categories. Although it generally shows lower precision values, its recall scores are more consistent. It excels in No-Impaired detection (0.76) and performs reasonably across other categories. This balanced approach suggests that the Advanced CNN might be more reliable for general screening, where missing cases may be more critical than occasional false positives.

The F1 Scores provide fascinating insights into the overall effectiveness of both models. The Basic CNN achieves higher F1 Scores in Very Mild Impairment (0.59) and Mild Impairment (0.53), while the Advanced CNN performs better in No Impairment cases (0.67). This pattern suggests that each model has developed different strengths in identifying specific impairment patterns.

From a clinical perspective, these results present essential implications. The Basic CNN's high precision makes it potentially more suitable for confirmatory testing, where false positives must be minimised. The Advanced CNN's more balanced metrics might make it more appropriate for initial screening purposes, where missing cases (false negatives) could have serious consequences.

Visual Analysis of Misclassified Cases:

1. **Pattern of Misclassification:**

The displayed brain scan images show three cases in which the model incorrectly classified Mild Impairment as Very Mild Impairment.
These images show subtle structural differences in brain tissue and ventricular shapes.
The consistent misclassification pattern suggests the model struggles to differentiate between mild and very mild cases.

2. **Misclassification Statistics:**

Mild Impairment: 88 misclassified cases (highest misclassification rate)
No Impairment: 397 misclassified cases (the most significant number)
Very Mild Impairment: 67 misclassified cases
Moderate Impairment: 10 misclassified cases (lowest number)

3. Clinical Implications:

The model shows particular difficulty distinguishing between adjacent impairment levels

There's a concerning high number of misclassifications for No Impairment cases

The relatively low misclassification rate for Moderate Impairment might be due to more distinct imaging features

The consistent confusion between Mild and Very Mild cases suggests a need for better feature extraction in subtle cases

0. Technical Insights:

The model appears to be sensitive to subtle variations in brain structure

There might be insufficient distinctive features learned for separating mild from very mild cases

The high number of No Impairment misclassifications suggests potential bias in the model

The visualisation reveals the challenge of dealing with gradient conditions rather than distinct categories

The visualisation effectively communicates these differences through side-by-side bar comparisons, with blue bars representing the Basic CNN and orange bars showing the Advanced CNN's performance. The clear distinction in performance patterns across classes highlights the importance of considering per-class accuracy rather than just overall model performance, particularly in medical applications where accurate diagnosis across all impairment levels is crucial.

The code and resulting visualisation analyse class-wise accuracy for the brain impairment classification model. The implementation uses Matplotlib to create a bar chart visualisation of accuracy metrics across different impairment classes. The code effectively utilises figure sizing (fig size=(10, 6)) and appropriate styling, with a sky-blue colour scheme, for a clear visual representation.

The accuracy distribution across classes reveals significant variations in model performance:

1. No Impairment shows the highest accuracy at 76.25%, indicating the model is most effective at identifying cases without cognitive impairment. This strong performance suggests the model has successfully learned the distinguishing features of healthy brain scans.
2. Moderate Impairment achieves a reasonable accuracy of 58.33%, positioning it as the second-best performing category. This suggests the model has moderate success in identifying the distinctive patterns associated with this level of impairment.
3. Mild Impairment shows lower performance with 47.49% accuracy, highlighting the challenge in detecting subtle early-stage impairment patterns. This reduced accuracy likely reflects the difficulty in distinguishing mild cases from normal and very mild cases.
4. Very Mild Impairment demonstrates the lowest accuracy at 12.50%, indicating a significant challenge in correctly identifying cases in this category. This poor performance suggests difficulties in distinguishing subtle variations in brain patterns characteristic of mild impairment.

4. Future Directions

Machine learning (ML) integration into Alzheimer's disease (AD) research has already demonstrated significant promise, but future advancements are necessary to maximise its potential. This section highlights key areas for innovation and collaboration to improve further early diagnosis, disease progression prediction, and personalised treatment.

4.1 Integration of ML with Emerging Technologies

Genomics and Proteomics

Advancements in genomics and proteomics offer a wealth of information about the molecular mechanisms underlying AD. By integrating ML with these technologies:

- Researchers can analyse complex datasets to uncover new biomarkers, such as genetic variations (e.g., APOE genotypes) and protein expression profiles (e.g., amyloid-beta and tau levels).
- ML algorithms can facilitate the identification of genetic risk factors and predict individual susceptibility to AD with greater accuracy.

Integrating ML with CRISPR-based gene-editing technologies may also enable researchers to model AD in vitro, improving our understanding of disease pathology and testing targeted therapies.

4.2 Development of Multi-Modal ML Models

Multi-modal ML models that combine diverse data types, such as neuroimaging, clinical data, genomics, and patient-generated health data, have the potential to provide a holistic view of AD.

- **Enhanced diagnostic accuracy**: Combining neuroimaging data with biomarker levels and genetic profiles allows for more precise AD detection in its earliest stages.
- **Comprehensive progression prediction**: Multi-modal models can incorporate longitudinal data to capture the temporal dynamics of disease progression.
- **Personalised medicine**: ML can generate treatment recommendations tailored to individual needs and responses by integrating patient-specific data from multiple sources.

The development of such models will require advanced ML architectures, including deep learning networks capable of handling heterogeneous data and techniques for fusing data from disparate sources.

4.3 Collaborative Efforts Across Disciplines

The complexity of AD research necessitates collaboration between researchers, clinicians, and data scientists. Key initiatives include:

- **Interdisciplinary research teams**: Facilitating knowledge exchange among neuroscience, bioinformatics, and ML experts to design innovative solutions.
- **Standardisation of datasets**: Collaborative efforts to create standardised, high-quality datasets that can be shared across institutions to enhance reproducibility and scalability.
- **Translational research**: Bridging the gap between laboratory research and clinical practice to ensure that ML advancements benefit patients directly.

Encouraging collaboration between academia, industry, and government organisations can accelerate progress in AD research and treatment development.

4.4 Scaling ML Solutions to Underserved Regions

Global disparities in access to healthcare and advanced diagnostic tools persist. ML solutions must benefit diverse populations, including those in underserved regions. Strategies include:

- **Low-cost diagnostic tools**: Leveraging ML algorithms to develop affordable and non-invasive screening methods, such as blood-based biomarker tests or smartphone-based cognitive assessments.
- **Localised training datasets**: Collecting and incorporating data from underrepresented populations to ensure inclusive and generalisable models.
- **Cloud-based platforms**: Deploying ML-powered tools via cloud infrastructure to enable remote access for clinicians and patients in resource-limited settings.

Addressing these disparities is essential to make ML-driven advancements in AD research and care accessible to all.

5. Conclusion

Alzheimer's disease (AD) remains one of the most pressing global health challenges, with its rising prevalence placing immense socio-economic and emotional burdens on patients, families, and healthcare systems. Machine learning (ML) has emerged as a transformative tool in AD research, offering unprecedented opportunities to advance early diagnosis, prognosis, and personalised treatment strategies.

By analysing large, complex datasets, ML algorithms have significantly enhanced our understanding of AD biomarkers, including neuroimaging, genomics, proteomics, and patient-generated health data. Through ML, researchers have developed accurate models to identify early-stage AD, often before clinical symptoms appear. These advancements allow timely intervention, which is critical for slowing disease progression and preserving cognitive function.

In prognosis, ML has contributed to predicting AD trajectory by integrating diverse data types, including longitudinal imaging, biomarker dynamics, and patient demographics. These predictive models enable clinicians to anticipate disease progression, facilitating more informed patient care and resource allocation decisions. Furthermore, ML-driven tools for multi-modal data integration pave the way for more robust and comprehensive assessments of disease progression.

ML has also benefited the field of personalised medicine. By tailoring treatment plans to an individual's unique biomarker profile, genetic predispositions, and lifestyle factors, ML has brought precision medicine closer to reality. This approach improves therapeutic outcomes and reduces the risks of one-size-fits-all treatment protocols.

Despite its immense potential, the application of ML in AD research is not without challenges, including the need for high-quality, standardised datasets and the ethical complexities of using sensitive patient data. Nonetheless, as interdisciplinary collaborations grow and technological innovations continue, ML stands as a cornerstone in reshaping the landscape of AD research, offering hope for better management and, ultimately, prevention of this devastating disease.

The future of ML in Alzheimer's research lies in harnessing emerging technologies, fostering interdisciplinary collaborations, and addressing the diverse needs of global populations. ML can transform the landscape of AD diagnosis, prognosis, and treatment by integrating genomics, proteomics, multi-modal data, and scaling solutions to underserved regions.

References

1. Alzheimer's Association. (2023). *2023 Alzheimer's Disease Facts and Figures*. Retrieved from https://www.alz.org
2. Ashrafian, H., Teare, H. J., & Bossuyt, P. M. (2020). Artificial intelligence and big data in public health: Opportunities and challenges. *Nature Medicine, 26*(5), 527–532. https://doi.org/10.1038/s41591-020-0832-4
3. Bron, E. E., Smits, M., Niessen, W. J., & Klein, S. (2015). Feature selection based on SVM classification for computerised dementia diagnosis. *Frontiers in Ageing Neuroscience, 7*, 20. https://doi.org/10.3389/fnagi.2015.00020
4. Cheng, B., Liu, M., Zhang, D., & Shen, D. (2017). Robust multi-label transfer feature learning for early diagnosis of Alzheimer's disease. *Brain Imaging and Behaviour, 11*(4), 1217–1230. https://doi.org/10.1007/s11682-016-9600-3
5. Esteva, A., Robicquet, A., Ramsundar, B., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z
6. Initiative for Open Access Series of Imaging Studies (OASIS). (n.d.). *OASIS Brain Dataset*. Retrieved from https://www.oasis-brains.org
7. Jiang, J., Liu, T., Zhu, W., & Shen, D. (2020). Machine learning methods for detecting Alzheimer's disease. *Computational and Structural Biotechnology Journal, 18*, 1627–1640. https://doi.org/10.1016/j.csbj.2020.06.016
8. Kaggledatasets. (n.d.). Alzheimer's Disease Detection Dataset. Retrieved from https://www.kaggle.com
9. Livingston, G., Huntley, J., Sommerlad, A., & Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet, 396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6
10. Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big data application in biomedical research and health care: A literature review. *Biomedical Informatics Insights, 8*(Suppl. 1), 1–10. https://doi.org/10.4137/BII.S31559
11. Minhas, S., Mehmood, R. M., & Cheema, M. A. (2022). Automated prediction of Alzheimer's disease using brain MRI images: A machine learning approach. *Journal of Neuroscience Methods, 375*, 109622. https://doi.org/10.1016/j.jneumeth.2022.109622

12. National Institutes of Health (NIH). (2021). *NIH Alzheimer's Disease Research Centres*. Retrieved from https://www.nia.nih.gov

13. Petersen, R. C., & Jack, C. R. (2022). Neuroimaging in Alzheimer's disease: Progress and promise. *Neuron, 110*(6), 952–970. https://doi.org/10.1016/j.neuron.2022.01.021

14. Quer, G., Arnaout, R., Henne, M., & Marcus, G. M. (2021). Wearable devices and the promise of artificial intelligence in healthcare. *Nature Biotechnology, 39*(9), 956–960. https://doi.org/10.1038/s41587-021-01060-3

15. Zhang, Y., Wang, S., Wang, L., & Yuan, Y. (2019). Deep learning in Alzheimer's disease: Algorithms, applications, and future directions. *Artificial Intelligence in Medicine, 101*, 101740. https://doi.org/10.1016/j.artmed.2019.101740

16. World Health Organisation (WHO). (2022). *Global Status Report on the Public Health Response to Dementia*. Retrieved from https://www.who.int