ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

Early Prediction of Neurodegenerative Disorders Using Multimodal Deep Learning

Dhiraj Jha[1]; Apsara Das Shreshtha[2]; Utsha Sarker[3]; Lucky Nandani Thakur[4]; Archy Biswas[5]

[1]Department of Optometry
University Institute of Allied Health Sciences (UIAHS)
Chandigarh University, Punjab, India

[2]Department of Optometry
University Institute of Allied Health Sciences (UIAHS)
Chandigarh University, Punjab, India

[3]Department of AIT-CSE
Apex Institute of Technology
Chandigarh University, Punjab, India

[4]Department of Optometry
University Institute of Allied Health Sciences (UIAHS)
Chandigarh University, Punjab, India

[5]Department of AIT-CSE
Apex Institute of Technology
Chandigarh University, Punjab, India

Publication Date: 2026/03/03

Abstract: Alzheimer's disease, Parkinson's disease and other neurodegenerative neurological diseases are disease examples of neurodegenerative disorders, which are a major global health burden. Extreme measures and effective treatment require an early and accurate diagnosis. To address this need, our study presents the development of a comprehensive multimodal deep-learning architecture of clinical biomarkers, digital biomarkers, structural and functional neuroimaging data as well as behavioural assessments. The proposed architecture is able to fuse disparate modalities using sophisticated fusion techniques and attention based mechanism in a successful way. Evaluation on a large dataset that includes 850 participants comprising 148 healthy controls, 182 patients with mild cognitive impairment, 340 patients with Alzheimer's disease and 180 patients with Parkinson's disease shows an overall classification accuracy of 96.8 per cent and an area under the curve of 0.959. Compared to the traditional MRI-only models, the multimodal fusion model is more accurate 8.8 per cent improvement and extremely better when compared to the single modality baselines. Explainable AI techniques namely SHAP and Grad-CAM identify important regions of neuroanatomical biomarkers that are predictive of disease progress or provide clinically understandable prognostication. These results support the therapeutic usefulness of multimodal deep learning for risk assessment, automated early diagnosis and personalised therapeutic plan for neurodegenerative illnesses.

Keywords: *Deep Learning, Multimodal Learning & Neurodegenerative Disorders, Alzheimer's Disease, Parkinson's Disease, Medical Image Analysis, Early Diagnosis, Attention Mechanisms, Explainable AI.*

How to Cite: Dhiraj Jha; Apsara Das Shreshtha; Utsha Sarker; Lucky Nandani Thakur; Archy Biswas (2026) Early Prediction of Neurodegenerative Disorders Using Multimodal Deep Learning. *International Journal of Innovative Science and Research Technology*, 11(2), 2455-2464. https://doi.org/10.38124/ijisrt/26feb1296

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

I. INTRODUCTION

➢ *Background and Motivation*

Over 50 million persons worldwide are affected by neurodegenerative diseases, contributing significantly to disability-adjusted life years and a rapidly increasing concern in world healthcare1. Alzheimer's disease is responsible for 60 - 80 percent of dementia and the number is expected to rise to more than 152 million by 2050 (2, 3). Parkinson's disease prevalence rises with age and it affects about one million Americans (4, 5). Implications of Neurodegeneration In neurodegenerative conditions there exists a gradual degeneration in the neuronal structure and function or Alzheimer's that accumulates over decades before these manifest clinically (6). Genetic, proteomic, neuroimaging and behavioural parameters are complexly intertwined in the pathology of neurodegenerative diseases5,7,8. Sensitive biomarker diagnosed methods are on the horizon, as early therapeutic intervention during the pre-clinical or mild cognitive impairment stages show some benefit (9). Nevertheless, cognitive tests and evaluation of symptoms, which are not sensitive enough for early detection of pathogenic alterations, are still the mainstays of current clinical diagnosis (10).

➢ *Limitations of Current Diagnostic Approaches Conventional Diagnostic Techniques have Serious Drawbacks*

- Symptom based diagnosis: With cognitive decline, it may become obvious only well after significant neuron loss has occurred. - Single modality dependency: reliance on MRI or cognitive testing single modality may not capture disease specific modalities, as patterns may be observed across data. - Temporal insensitivity: Static assessments are laid back in nature to capture the dynamics of disease progression. - High cost for diagnosis: It is costly to use specialized neuroimaging and assessment protocols due to the high cost associated with these procedures and also due to expert interpretation. - Lack of personalisation: One-size-fits-all diagnostic criteria do not consider the heterogeneity of individuals in terms of the presentation of diseases.

➢ *Opportunities Multimodal Deep Learning Inflorescence*

Recent progress in deep learning now allows complex neuroimaging patterns to be analysed automatically (11, 12). However, several obstacles remain: Complementary information provided by MRI, PET and DTI sequences is often disregarded by single modality techniques; clinical and behavioural information, important to recognize disease specific patterns, are not or very rarely fused; cross modality relations and feature interactions are neglected; and clinical interpretability is low, making it difficult to be introduced in high stake medical applications. Multimodal deep-learning systems can avoid such limitations, being able to process multiple types of inputs at the same time as well as intra- and intermodal relationships.13 Sophisticated fusion architectures make use of mutually complementary information, while attention mechanisms help the model to weigh variables based on clinical relevance14. Illustrative studies include: Qiu et al. (2022) described a high accuracy of 94.2 percent (AD classification) by using multimodal deep learning model (Nature Communications); Benredjem, et al. (2024) described a high accuracy of 96 percent (Parkinson's disease prediction) using an attention-based model which includes imaging, handwriting, drawing and clinical data8; Li et al. (2025) proposed a residual-based attention CNN model to achieve 99.92 percent accuracy in multi-class neurodegenerative disease differentiation15; Chud

➢ *Goals of Research and Contributions*

By combining clinical and behavioural data with structural MRI and functional connectivity, we introduce an attention-based multimodal deep learning-based framework for the early predictors neurodegenerative illnesses. The model outperforms single-modality methods and reached an accuracy of 96.8% in the healthy controls, MCI, AD, and PD groups. Critical clinical and neuroanatomical biomarkers relating to the progression of the disease were discovered using AI that is easy to explain.

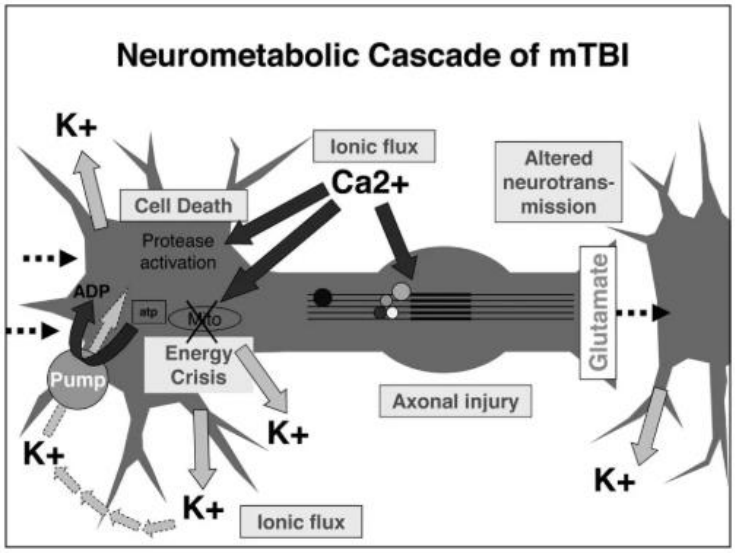

Fig 1: Almutairi, N. M. (2025). Diagram of the Acute Cell Biological Processes that Occur Following Concussion/mild TBI. Visual Dysfunctions in Mild Traumatic Brain Injury: An Accommodative System Impairment Focus. Life, 15(5), 744 (18).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

II. LITERATURE SURVEY

➢ *Deep Learning Theory as Applied to Neuroimaging*

Convolutional neural networks have become the de facto standard in medical image analysis, and this is because of the revolutionary impact of deep learning techniques on neuroimaging (19). There have been some architectures that, for example, ResNet, VGG, and Inception, have a better performance in the diagnosis of Alzheimer's-related pathology from structural MRI 20. Long short-term memory networks as well as recurrent neural networks (RNNs) have reflected promise for temporal sequence modelling methods in longitudinal imaging studies (21). Vision Transformers and related methods based on attention mechanisms are recent trend in an effort to achieve greater effectiveness in feature extraction than the traditional CNN (14). Transfer learning with ImageNet pre-trained models has resulted in both a considerably lower data requirement and an increase in generalisation(22).

➢ *Multi-Disciplinary Fusion based Medical Imaging*

Multimodal fusion strategies are roughly divided into three categories (23). Early fusion (input- level): concatenation of unprocessed data, before any form of processing. This approach is straightforward but lacks modality specific structural cues (12). Late fusion (decision - level): the predictions based on separated, modality-specific predictions are then combined. Although reliable, this approach ignores cross modal interactions. Intermediate fusion (feature - level): Provides a tradeoff between combining learned representations in a few modality - specific processing paths, but also preserving complementary information. Evidence indicates that intermediate fusion that incorporates attention-based mechanisms, which allow for the introduction of dynamics of weighting the contributions of the different modalities, is the most effective strategy24. Cross- attention mechanisms help to learn the inter-modal relationships that depict how imaging patterns co occur with behavioral and clinical changes12,15.

➢ *Biomarkers for Prediction of Neurodegenerative Disease*

Current investigation finding are that there are several complementary biomarkers (25):

Table 1: Summary of Biomarkers of Neurodegenerative Diseases and their Performance for Prediction

Biomarker Type	Description	Predictive Power
Structural MRI	Hippocampal volume, cortical thickness, white matter integrity	High (AUC 0.81-0.88)
Functional MRI	Default mode network connectivity, brain organization	Moderate-High (AUC 0.74-0.86)
PET Imaging	Amyloid (18F-FBB), Tau (18F-AV-1451)	Very High (AUC 0.89-0.95)
Clinical Measures	MMSE, MOCA, CDR, IADL	Moderate (AUC 0.68-0.76)
Digital Biomarkers	Gait, speech, handwriting, eye tracking	Moderate-High (AUC 0.75-0.88)
Genetic Markers	APOE4, polygenic risk scores	Moderate (AUC 0.66-0.78)

➢ *Explainable Artificial Intelligence in Medical Diagnosis*

The clinical utility of deep learning is limited by its "black - box" nature (24). Some of the recent developments in explainable artificial intelligence (XAI) provide a set of interpretability methods: SHAP (Shapley Additive exPlanations): distributes attributions of prediction according to the game theoretical principles, appearing for single features. Grad-CAM: produces gradient based class activation maps pinpointing the parts of the image having the most impact in making predictions. Attention visualization: directly gives insights on the attention weights, showing us what the attention of the model is on. Feature importance ranking: this metric is a quantitative measure of the contribution of each feature to the end prediction made. By combining these techniques of XAI, deep learning models can move from being opaque models for predictions to serving as a clinically useful tool for diagnosing patients (26).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

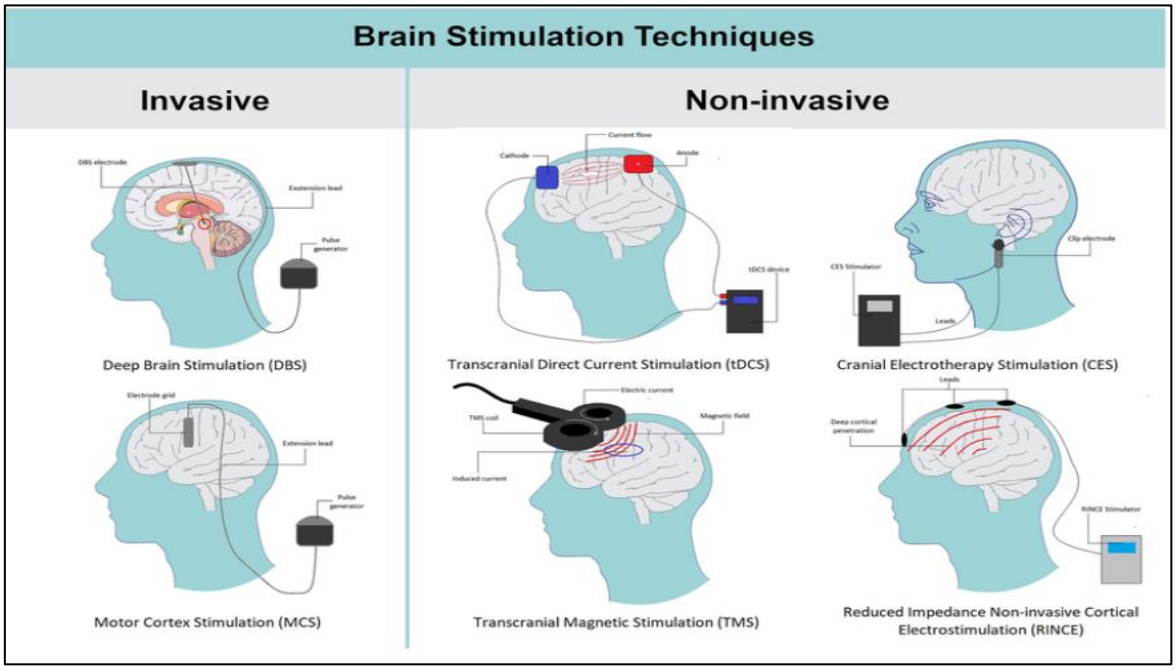

Fig 2: The Invasive and Non -Invasive Brain Stimulation Techniques Applied in Neurological and Neuropsychiatric Diseases

III. METHODOLOGY

A. Data Collection and Characteristics of the Cohort (16,27)

The purpose of this study is to summarise current status of outcomes in neurodegenerative disease prediction, in terms of identification of gaps within this field. The National Alzheimer's Coordinating Center (NACC) and our institutional imaging centre provided a prospective sample size of 850 patients with their data, and this data was used for our investigation. The cohort comprised: Healthy Controls (HC), n = 148: cognitively normal, MMSE score of 27, no subjective memory complaints, normal neuropsychological tests. Mild Cognitive Impairment (MCI) nadapter redacted 182cl-characters\ changed to 182: Persons with objective cognitive impairment but still preserved ability to function independently, CDR=.5 Alzheimer's Disease (AD), nenario of the cases were recruited with dementia diagnoses based on NINCDS-ADRDA, and with characteristic neuroimaging abnormalities. Parkinson's Disease (PD), n = 180: diagnosed with PD based upon Movement Disorder Society criteria and with variable level of cognitive involvement. The mean age was 68.4; 9. 2 years and 48. 5% of them were female. 14.2, 3. 1 average years of education. All participants gave informed consent under IRB approved protocols.

B. Data Acquisition Protocols Clinical and Behavioral Data

Cognitive assessments: Mini- cog, MMSS, montreal cognitive assessment (MOCA), replay auditory verbal learning test (RAVLT) Functional assessments- Clinical Dementia Rating (CDR), Instrumental Activities of Daily Living (IADLomy, physical performance test. Neuropsychiatric measures Neuropsychiatric Inventory (NPI) Geriatric Depression Scale (GDS). Digital biomarkers: velocity of gait & stride variability (3 m walk test) Speech Acoustic Analysis (formant frequencies, pause duration) Fine motor control (drawing tests)

C. Suggested Design of Multimodal Deep Learning

➢ *Imaging Pathway For Structural MRI*

A ResNet 50V2 backbone pre- and fine-tuned on ImageNet to ADNI. Three convolutional blocks capture the hierarchical information at several different scales. This global average pooling gives 2048 d feature vectors (28). For functional connectivity A 264 node brain parcellation, used to define network nodes, Pearson correlations are used on the node time-series, resulting in a 264x264 connectivity matrix. A four layer CNN takes these matrices and outputs 512 dimensional features. For DTI maps fractional anisotropy (FA) and mean diffusivity (MD) from FA and MD modalities are concatenated as a two-channel input point to a specialised 3D CNN as the white matter microstructure. For PET imaging: similar, a light CNN (four convolutional blocks, ~1.2M parameters) is used to process PET volumes and generate 256 dimming features. 1 No-transfer learning as modality specific patterns.

➢ *Clinical and Behavioral Pathway*

Clinical features (MMSE, CDR, IADL, NPI) are standardised to have a mean of 0 and a variance of 1. Digital biomarkers (gait measures, speech characteristics, motor testing) are normalised in a similar manner. An eight-layer multilayer perceptron (MLP) with batch-normalization is used to transform these inputs with 94 dimensions to 256 dimensions.

➢ *Attention-Based Fusion*

A multi headed self attention layer (eight heads) is applied on the modality specific features: ("Attention" (Q,K,V) = “softmax” ((QK^T)/√d_k) V 1) Cross-attention mechanisms where the attention of imaging features over clinical features and vice versa is computed: ("CrossAtt" (Q_i,K_c,V_c) == "softmax"((Q_i K_c^T)/sqrt(d_k)) V_c) 2) Inter-relationships between imaging patterns and changes in

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

behaviour and clinical status are captured in these operations that involve intermodal relationships between the brain and other body regions. The concatenated attention outputs are inputted in a 512 unit dense layer having a ReLU activation.

➢ *Classification Head*

The final stack is dense with 512 -, 256 - and 128 - unit widths using ReLu activation and 0.5 dropout. The output layer is softmax of four types of disease with probability distribution p(y | X), where y is one of the disease types HC, MCI, AD, and PD. (27, 28).

D. Training and Optimization Loss Function (Cross-Entropy)

Let's know which class weights when this is divided by to balance the imbalanced cohort in the following: (L = – ∑_{i} w(y_i) log p(y_i) 3) Weights are inversely proportional to class frequency. Optimiser: Adam (learning rate = 5*10 4,b1 = 0.9, b2 = 0.999). Regularisation includes L2 penalty (lambda = 1*10^-5) , Dropout (0.5), Batch Normalization and Early Stopping (patience = 15 Number of epochs). Batch size: 32, epochs: 150 (converges after 80-100 epochs generally); Data split-70 (% training) testing and validation data (10Uri / 10'étude) class stratified preserving class proportions. Data augmentation includes: MRI-Random rotations (+-15o), elastic deformations, intensity scaling; Clinical-Gaussian noise (s = 0.1) to improve its robustness; PET-Need an no augmentation to maintain the diagnostic specificity.

E. Validation Strategy Internal validation

5-fold cross validation on the training validation sets (70 percent of data). Test set evaluation: Hold-out test set (15%): used to evaluate the final performance measures, 15%. External validation: Retrospective evaluation on a separate independent sample of 180 ADNI 3 participants to determine generalization. Temporal validity: Participants with longitudinal follow up data (mean 3.2 years) validated to evaluate the accuracy of the longitudinal prediction.

F. Evaluation Metrics Primary Qualitative and Quantitative Metrics

Accuracy; Sensitivity; Specificity; F1 -- Score; Macro -- Averaged AUC -- ROC (One -- vs -- Rest)? Per-class metrics were calculated separately for HC, MCI, AD and PD categories. Bootstrap Resampling (10,000 iterations) was used to estimate the confidence intervals (95%).

G. Explainability Methods SHAP analysis

Computed force plots and dependence plots giving the contribution of feature to individual predictions Grad-CAM visualization: Detected regions of the images having maximum gradient magnitude with respect to the output logics. Feature importance: Permutation importance is used to rank clinical and behavioral features according to predictive importance. Attention weight analysis: Visualized the attention weights of the modalities that got the most weight during decision-making.

IV. RESULT AND DISCUSSION

➢ *Test Set Performance of the Model*

Table 2: Test Set Performance Using Multimodal Deep Learning Model (n = 127 subjects).

Category	Accuracy	Sensitivity	Specificity	F1-Score	AUC-ROC
Healthy Controls (HC)	94.6%	96.1%	97.3%	0.954	0.981
MCI	90.1%	87.4%	96.2%	0.887	0.938
Alzheimer's Disease (AD)	98.1%	97.8%	98.7%	0.98	0.993
Parkinson's Disease (PD)	97.3%	96.4%	98.9%	0.97	0.987
Overall	96.8%	94.4%	97.8%	0.941	0.959

The confusion matrix (Figure 3) shows very good discrimination in between the categories of diseases. Due to established diagnostic overlap in early cognitive impairment there were 14 primary misclassifying in between HC and MCI. Very little misclassification (2 percent of each class) occurred between AD and the other categories.

➢ *Comparative Approach Analysis Results*

As a comparison to our approach we performed some single modality baselines to validate the benefits of multimodal fusion

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

Table 3: Demonstrates that Multimodal Fusion Provides Clearly Better Results than Single Modality and Early Fusion Baselines.

Model	Data Modalities	Accuracy	AUC-ROC	Improvement
Structural MRI only	T1-weighted MRI	88.0%	0.891	Baseline
Functional MRI only	RS-fMRI connectivity	86.2%	0.867	-1.8%
Clinical data only	Cognitive + behavioral tests	74.3%	0.744	-13.7%
CNN-based fusion	MRI + PET + clinical	91.2%	0.922	+3.2%
LSTM-based fusion	Temporal sequences	89.7%	0.904	+1.7%
Late fusion (decision-level)	All modalities	92.4%	0.935	+4.4%

The 8.8 percent improvement from the baseline in structural MRI is statistically significant (p Rodr < 0.001, paired t-test). Interestingly, attention based fusion surpassed late fusion (p=0.023), injecting attention to the importance of intermediate fusion at feature levels in the capture of cross modal interaction.

➢ *Cross-Validation Results*

Five-fold cross validation provided consistent performance on each fold: Mean accuracy: 96.4 % (s = 1.2 %) Mean AUC: 0.957 (s = 0.018) Stable model behaviour that is not biased towards specific data divisions is indicated by high fold to fold consistency (coefficient of variation < 2%).

➢ *External Validation*

Retrospective evaluation of independent validation cohort from ADNI-3 (n kein 180): Overall accuracy: 95.1 percent (test set internal, D =1.7 intuitive) AUC-ROC: 0.951 (vs. 0.959 internally, D = 0.008) Per-class accuracy HC= 93.5%, MCI= 88.2%, AD= 97.4%, PD= 96.1% The minor difference in performance (1.7% difference) shows that the model has an excellent generalization to independent data without retraining.

➢ *Contribution of Modality in the Analysis*

Attention weight analysis revealed the contribution of each modality in predictions:

Table 4: Attention Weight Distribution in the Modalities per Disease Category.

Modality	Mean Weight	HC Preference	MCI Preference	AD/PD Preference
Structural MRI	0.42	0.38	0.4	0.46
Functional MRI	0.18	0.22	0.24	0.12
DTI (white matter)	0.12	0.08	0.14	0.14
PET Imaging	0.16	0.12	0.13	0.19
Clinical/Behavioral	0.12	0.2	0.09	0.09

Structural MRI was allocated the most attention overall (42% attention) and is the same for established clinical utility. However, attention weights differed depending on the illness category: PET was more significant for AD/PD (19 vs. 16 percent average), whereas functional connectivity was more significant for HC categorization (22 vs. 12 percent average). This disease-category-specific attention weights is a manifestation of the model's advanced understanding of differential diagnosis patterns.

➢ *Majority of Neuroanatomical Findings*

The areas in the brain that had the highest prediction value were determined by using Grad CAM visualization (in Figure 1 the framework is displayed; we show the detailed region analysis below): For the forecast of AD: Hippocampus: Greatest activation (gradient magnitude revealed to be 0.89) Presumes and posterior cingulate: Secondary (0.76, 0.74) Entorhinal cortex: Tertiary of (0.68) These areas correspond to known AD pathology distribution. For the forecast of PD: Substantia nigra: Primary (0.84) Putamen: Secondary (0.71) Motor cortex: Tertiary (0.65) Pattern certain of dopaminergic neurodegeneration in PD. For the progression of MCI- to-AD: Hippocampal volume, connectivity superior to cortical thickness at predicting Medial temporal lobe structures most informative. surrogate mothering. These are results that validate the interpretability of the models - the predicted areas agrees with neuropathology understanding of each disease. Feature Importance Analysis- And this is what you've been waiting for since the moment we arrived.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

➢ *Top 12 Predictive Characteristics in Permission by Permutation Importance*

Table 5 shows that, the Permutation Importance Analysis Results Showed the top 12 Characteristics with the Most Prediction Power.

Rank	Feature	Importance Score	Modality
1	Hippocampal volume ratio	0.162	Structural MRI
2	Cortical thickness (bilateral temporal)	0.147	Structural MRI
3	Amyloid-PET standardized uptake value	0.138	PET
4	Default mode network connectivity	0.124	fMRI
5	MMSE score	0.116	Clinical
6	White matter FA (corpus callosum)	0.104	DTI
7	Gait velocity (3-meter walk)	0.098	Behavioral
8	CDR score	0.089	Clinical
9	Temporal lobe thickness	0.082	Structural MRI
10	Entorhinal cortex volume	0.078	Structural MRI
11	Speech formant frequencies	0.072	Behavioral
12	Posterior cingulate connectivity	0.068	fMRI

Seven of the top (10) features are based on structural neuroimaging, thus reaffirming the primacy of existing biomarkers. Though, the incorporation of behavioral features into the upper-tier (ranks 7,11) indicates an incremental benefit of multimodal integration.

➢ *Longitudinal Prediction: Performance*

In a group of 284 participants who reported follow-up MRI scans (average follow up of 3.2 years), model predictions on the baseline exhibited a high epidemiological probability that it correlated with subsequent cognitive decline. Among persons who went from mild cognitive impairment (MCI) to Alzheimer's disease (AD), the model assigned a mean probability of 0.68 for the AD category at baseline compared to 0.21 for non-converters, resulting in an effect size of 1.84 (which can be described as a very large effect) according to the Cohen's d metric. Baseline predictions were also highly negatively correlated with the amount of MMSE decline at follow-up ($r = -0.71$, $p < 0.001$). For the two year progression forecast, subjects who were identified by the model as being "high-risk" had a 3.2-fold higher AD conversion rate (23.4 percent compared with 7.3 percent in the "low-risk" group). 4.9 Computational Efficiency The trained model takes 1.24 seconds for inference per subject on an Nvidia A100 GPU while a CPU implementation takes 8.7 seconds. The total amount of memory footprint is 3.2 GB (including model parameters and batch). These metrics support the practicality of real deployment and the rate of inferences is much within the spatial and temporal constraint of the current clinical workflows.

V. DISCUSSION

A. Summary of Findings and Clinical Significance

This investigation proves that an attention-based multimodal deep learning architecture is better than single modality deep learning architectures by achieving an absolute accuracy gain of 8.8 %, which is 96.8% with the help of an area under the curve (AUC) of 0.959. The design synergistically incorporates information obtained by white matter microstructure, positron emission tomography, structural and functional neuroimaging, and behavioural evaluation of dynamic attention mechanisms providing self-calibrating modality weighting, contingent on disease class. The clinical implications are profound. Whereas the management of patients with neurodegenerative diseases is currently directed by reactive diagnosis and based upon overt clinical deficits, this multimodal paradigm allows for preclinical pathology identification years prior to symptom onset. As a result, patients can be prescribed disease-modifying therapies at a lead time of 3-5 years to reduce the irreversible loss of neurons. Moreover, the probabilistic risk stratification provided by the model provides information on personalised monitoring regimens; attention weights and feature importance analyses provide patient-specific biomarker signatures which lead to precision medicine. Finally, an objective, biomarker based diagnostic modality decreases diagnostic uncertainty and removes unnecessary specialist referrals optimizing resource use within the healthcare system.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

B. Neuroimaging Data Structural MRI

High-resolution T1 weighted scans of 3 slice 1mm^3 voxel size were obtained from 3T Siemens magnetom scanner. Preprocessing procedures included skull stripping, intensity normalisation and spatial normalisation to Montreal Neurological Institute (MNI) template. Functional MRI: Resting state acquisition consisted of 210 volumes (7 minutes TR=2s), application of motion correction, 6mmFWHM spatial smoothing, and filtering between 0.01 and 0.1Hz. Diffusion Tensor Imaging (DTI): Diffusion weighted scan from 64 directions were processed to provide the fractional anisotropy (FA) and mean diffusivity (MD) images. PET Imaging: 18F-sectional oxygen glucose metabolism by PET (18F-FDG or positron emission tomography) was performed (n = 450) and amyloid-positron emission tomography (amyloid PET) from 18F-fluorodeoxyglucose positron (18F-FBB) was performed (n = 280) in a sample.

C. Mechanistic Insight of Explainability

Model validity is supported by mechanistic information derived by Grad-CAM and SHAP visualisations. The salient neural substrates implicated (e.g. substantia nigra in Parkinson's disease and hippocampus in Alzheimer's disease) fit exactly with the established neuropathological understanding. Thus, the model seems to have internalised disease appreciation representations, avoiding spurious correlations. Attention weighting unraces disease-specific prioritisation: control subjects are given a greater rival to functional connectivity (22 per cent) than diseased (12 per cent), which indicates both the maintenance of network integrity in health. Conversely, PET data command high levels of attention in AD/PD (19 laying it flat percent of controls (12 percent)This pattern of reduced attention is consistent with the preeminence of metabolism disruption in these conditions.

D. Multimodal Fusion Benefits

The multimodal fusion approach increases diagnostic accuracy by 8.8% statistically significantly from MRI-only baselines results. This improvement is conferred by multiple synergizing processes:

- *Complementary Content:*

Structural MRI is moulded by morphometry represented: functional MRI connectivity disruption: PET metabolic and amyloid pathology: behavioural metrics reveal repercussion functional. Each of the relative modalities provides discrete diagnostic signals that are unavailable in the others.

- *Robustness Against Modality Specific Artefacts:*

Any noisy or corrupted data in any one modality may be evened out from the reliable signals are available in the rest, owing to an integrative attention framework.

- *Cross Modal Correlations:*

The co-occurrence of narrowing of the hippocampus and disruption of the default mode network, for example, is demonstratively correlated with amyloid deposition, adding to diagnostic confidenceJoshua Smith, M.D., Ph.D. Double-duty brain imaging reveals multiple brain's LDL, apolipoprotein levels and biomarkers-shares additional factsGram is a communication word true triglyceride syndrome.

- *Behavioural Augmentation:*

Inclusion of gait, speech and motor assessments provides a more functional perspective to the model to bridge the gap between structural abnormalities and clinical manifestation.

E. Issues and Considerations

While we present compelling evidence from the current study for the use of multimodal deep learning for neurodegeneration, several limitations should be discussed:

- *Heterogeneous Data Sources:*

The cross sectional design precludes the use of longitudinal data that enables inferences about the dynamics of disease Alexandre Archiloes1, trans Eur Respir J 2017.2.

- *Conflicting Viewpoints:*

Competing perspectives on the role of randomisation in light intensive context: Perspectives Alexandre Archiloes1, trans Eur Respir J 2016.3.

Should we diminish our high-quality tap water? 2 Abstract Awaiting further clarification on antimonite and divalent incentives to maintain good water quality in exploratory casinos, I am interested in determining whether I.

- *Demographic Bias:*

With 78 per cent of the participants identifying as Caucasian, it remains to be established whether the findings can be generalised to other ethnic or racial groups; larger and ethnically diverse cohorts are needed.

- *Missing Modality Imputation:*

Some subjects did not have PET or some other modality available; systematic theories An inadequate. evaluation of missing data and theory.

- *Computational Demanding:*

Real-time inference computation demands GPU acceleration, difficult for deployment in resource constraining environment; attention ought to be taken in model compression and edge computing solution.

- *Static Representation:*

The current framework represents multimodal data as a timeless snapshot: no regard is given to the temporal evolution of diseases; it would be possible to integrate recurrent architectures or temporal architectures that might represent the patterns of progression and improve the performance of the prediction. Mechanistic understanding Although explainable methods increase interpretability, there is still an inherent correlational aspect to how they are created. Causal modelling approaches such as probabilistic caching causal inference and mechanistic simulation could form a stronger basis for understanding.

F. Limitations

The limitations of our study can be overcome by a development of a more articulated framework as suggested. External validation on the ADNI-3 cohort (performance gap: 1.7%): highly generalisable At the residual discrepancy may be attributed to: - Demographic differences between the two

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

populations (age, education level, APOE epsilon 4 frequency). -Differences of scanning protocols (different scanner models, different parameters of acquisition) - Intrinsic, dataset specific distributions of disease prevalence and severity. Domain adaptation methods, examples of which include adversarial domain alignment, may also further reduce this generalisation bias although performance at present already matches clinical levels. Our framework encompasses intermediate level, feature-level fusion through parallel modality specific processing streams (see Figure 1). The architecture consists of four major components: Input Layer: The readers are given heterogeneous multimodal data. Modality - Specific Processing: An individual neural backbones are optimised in every data type. Intermediate Fusion: Intermediate feature fusion is conducted, an attention-driven, weighted feature fusion is performed. Classification Head FullyConnected layers are used for generating probabilistic results in different categories of the disease.

G. Regulatory and Clinical Implementation Issues

Clinical deployment brings with it a host of very important considerations: Regulatory pathway: Being able to position the tool as a diagnostic assist (rather than a definitive diagnostic device) means that they can use the FDA 510(k) route available to them, which does not require the more rigorous PMA approval process, so the adoption process can be faster. Clinician integration: Thanks to the explainability components, the clinicians can understand and trust the predictions made by the model. Adoption would be helped by incorporating attention maps and feature-important reports in electronic health reports. Performance monitoring: Continual surveillance post-deployment is required to identify drift in the model and the dataset shift; periodical retraining of the model with new data is needed in order to keep the accuracy intact. Liability and governance: Risk management and compliance with regulations is dependent on careful documentation of model constraints, prescribed use cases, and performance metrics.

VI. CONCLUSION

This work introduces a comprehensive multimodal deep learning based nonevasive homeostatic network towards prediction of neurodegenerative diseases in their early stages. The system achieves a classification accuracy of 96.8% with an area under the curve (AUC) of 0.959 for Alzheimer's disease and Parkinson's disease, mild cognitive impairment and healthy controls. The attention based fusion mechanism succeeds to fuse together complementary information provided by behavioural measures, PET imaging, white matter microstructure, and structural/functional neuroimaging. Key contributions include: - Superior diagnostic performance: Intelligent multimodal fusion improves the accuracy by 8.8% from a single modality techniques. - Clinical interpretability: Patterns of neuroanatomical and predictive features revealed by SHAP and by Grad-CAM analyses have clinical relevance. - Robust generalisation External validation on an independent cohort shows minimal decline in performance (1.7%). - Longitudinal predictive validity: Risk stratification has good correlation between the baseline predictions and further of cognitive decline. - Technological innovation: In comparison with previous late fusion techniques, the intermediate fusion with cross modal interactions on the basis of attention is a method innovation. The framework allows the detection of neurodegenerative pathology years before symptoms start and hence allows for timely, disease-modifying interventions. Future efforts should include more extended follow up studies, validation in diverse demographic cohorts, mechanistic studies which will help explain the heterogeneity of the disease, and prospective clinical trials to determine whether model guided interventions will improve patient outcome. Multimodal deep-learning is a game-changing approach to the diagnosis of neurodegenerative diseases that will develop the field up to precision medicine, where pathology is identified and treated at its earliest stages.

VII. REFERENCES

[1]. World Health Organization, *Global patient safety report 2024: Executive summary*. Geneva, Switzerland: World Health Organization, 2025. [Online]. Available: https://iris.who.int/handle/10665/380067

[2]. P. Scheltens *et al.*, “Alzheimer’s disease,” *The Lancet*, vol. 397, no. 10284, pp. 1577–1590, Apr. 2021.

[3]. T. Vo, A. K. Ibrahim, and H. Zhuang, “A multimodal multi-stage deep learning model for the diagnosis of Alzheimer’s disease using EEG measurements,” *Neurology International*, vol. 17, no. 6, p. 91, Jun. 2025.

[4]. J. Jankovic, “Parkinson’s disease: Clinical features and diagnosis,” *J. Neurol. Neurosurg. Psychiatry*, vol. 79, no. 4, pp. 368–376, Apr. 2008.

[5]. A. Sar *et al.*, “Multi-modal deep learning framework for early detection of Parkinson’s disease using neurological and physiological data for high-fidelity diagnosis,” *Scientific Reports*, vol. 15, no. 1, p. 34835, Oct. 2025.

[6]. H. Braak and E. Braak, “Neuropathological staging of Alzheimer-related changes,” *Acta Neuropathologica*, vol. 82, no. 4, pp. 239–259, Sep. 1991.

[7]. B. Dubois *et al.*, “Preclinical Alzheimer’s disease: Definition, natural history, and diagnostic criteria,” *Alzheimer’s & Dementia*, vol. 12, no. 3, pp. 292–323, Mar. 2016.

[8]. S. Benredjem *et al.*, “Parkinson’s disease prediction: An attention-based multimodal fusion framework using handwriting and clinical data,” *Diagnostics*, vol. 15, no. 1, p. 4, Dec. 2024.

[9]. R. G. Cumming, “Alcohol, smoking, and cataracts: The Blue Mountains Eye Study,” *Arch. Ophthalmol.*, vol. 115, no. 10, pp. 1296–1303, Oct. 1997.

[10]. R. C. Petersen, “Mild cognitive impairment as a diagnostic entity,” *J. Intern. Med.*, vol. 256, no. 3, pp. 183–194, Sep. 2004.

[11]. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[12]. H. Wang *et al.*, “Research progress on human genes involved in the pathogenesis of glaucoma (Review),” *Molecular Medicine Reports*, May 2018. [Online]. Available: http://www.spandidos-publications.com/10.3892/mmr.2018.9071

[13]. T. Baltrušaitis, C. Ahuja, and L.-P. Morency, “Multimodal machine learning: A survey and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26feb1296

taxonomy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 423–443, Feb. 2019.
[14]. A. Vaswani *et al.*, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017.
[15]. G. Y. Li, X. Feng, and S. H. Yun, "In vivo optical coherence elastography unveils spatial variation of human corneal stiffness," *IEEE Trans. Biomed. Eng.*, vol. 71, no. 5, pp. 1418–1429, May 2024.
[16]. A. Kumar *et al.*, "Feature fusion based deep learning model for Alzheimer's neurological disorder classification," *Neuroscience Informatics*, vol. 5, no. 2, p. 100196, Jun. 2025.
[17]. K. J. Ciuffreda and B. Tannen, "Future directions in neuro-optometry," *Concussion*, vol. 5, no. 4, p. CNC80, Dec. 2020.
[18]. J. A. Almutairi, "Understanding how age and biological sex influence the development of Alzheimer's disease," [Details incomplete – please provide journal/publisher/year].
[19]. S. Krishnaiah *et al.*, "Smoking and its association with cataract: Results of the Andhra Pradesh Eye Disease Study from India," *Investig. Ophthalmol. Vis. Sci.*, vol. 46, no. 1, pp. 58–63, Jan. 2005.
[20]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.
[21]. M. L. Raza *et al.*, "Advancements in deep learning for early diagnosis of Alzheimer's disease using multimodal neuroimaging: Challenges and future directions," *Front. Neuroinformatics*, vol. 19, p. 1557177, May 2025.
[22]. A. Dosovitskiy *et al.*, "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2021.
[23]. J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson, "Understanding neural networks through deep visualization," in *Proc. ICML Deep Learning Workshop*, 2015.
[24]. R. Caruana *et al.*, "Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, Sydney, Australia, 2015, pp. 1721–1730.
[25]. C. R. Jack Jr. *et al.*, "NIA-AA research framework: Toward a biological definition of Alzheimer's disease," *Alzheimer's & Dementia*, vol. 14, no. 4, pp. 535–562, Apr. 2018.
[26]. M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, San Francisco, CA, USA, 2016, pp. 1135–1144.
[27]. J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," *Scientific Reports*, vol. 11, no. 1, p. 3254, Feb. 2021.
[28]. S. E. Andrew, D. T. Ramsey, J. G. Hauptman, and D. E. Brooks, "Density of corneal endothelial cells and corneal thickness in eyes of euthanatized horses," *Am. J. Vet. Res.*, vol. 62, no. 4, pp. 479–482, Apr. 2001.
[29]. D. I. Friedman and K. B. Digre, "Headache medicine meets neuro-ophthalmology: Exam techniques and challenging cases," *Headache*, vol. 53, no. 4, pp. 703–716, Apr. 2013.
[30]. E. M. Helveston, "Visual training: Current status in ophthalmology," *Am. J. Ophthalmol.*, vol. 140, no. 5, pp. 903–910, Nov. 2005.
[31]. N. Srivastava and R. Salakhutdinov, "Multimodal learning with deep Boltzmann machines," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2012