

MULTIMODAL DEEP LEARNING FOR

Early Alzheimer's Disease Detection

Pre-Clinical Stage Risk Prediction Using MRI, PET, and Blood Biomarkers

Dataset: Alzheimer's Disease Neuroimaging Initiative (ADNI)

Architecture: CNN + Vision Transformer + Cross-Modal Attention Fusion

1. Alzheimer's Disease — An Overview

Alzheimer's disease (AD) is a progressive, irreversible neurodegenerative disorder and the most common cause of dementia worldwide. It is characterised by the gradual destruction of memory, thinking, behaviour, and ultimately the ability to carry out the most basic daily tasks. Unlike normal age-related cognitive decline, Alzheimer's involves the physical destruction of brain cells and the connections between them — a process that begins silently, years or even decades before any symptoms appear.

1.1 Global Burden and Epidemiology

Alzheimer's disease represents one of the most significant public health challenges of the 21st century. The scale of its impact is staggering and continues to grow as the global population ages.

Global Statistics

Over 55 million people worldwide are currently living with dementia, with Alzheimer's disease accounting for 60–70% of all cases — approximately 33–38 million individuals. By 2050, this number is projected to nearly triple to 153 million people, driven primarily by population ageing in low- and middle-income countries.

- A new case of dementia is diagnosed every 3 seconds worldwide
- In India, over 5.3 million people currently live with dementia, a number expected to reach 14 million by 2050
- Women are disproportionately affected — nearly two-thirds of all Alzheimer's patients are female
- The economic cost of dementia globally exceeds USD 1.3 trillion annually, including direct medical costs and informal caregiving
- Only 1 in 4 people with Alzheimer's disease has been formally diagnosed

1.2 Clinical Impact and Consequences

The consequences of Alzheimer's disease extend far beyond memory loss. The disease systematically destroys the brain's capacity for virtually every cognitive and functional domain, ultimately leaving patients entirely dependent on caregivers.

Cognitive consequences

- Progressive episodic memory loss — inability to retain new information
- Language deterioration — word-finding difficulties evolving to complete aphasia
- Executive dysfunction — loss of planning, reasoning, and problem-solving ability
- Visuospatial disorientation — inability to navigate familiar environments
- Loss of recognition — eventually unable to recognise family members

Neuropsychiatric consequences

- Depression and anxiety — present in up to 50% of patients
- Agitation, aggression, and sleep disturbances
- Psychosis, including hallucinations and delusions in late stages
- Apathy and social withdrawal

Systemic consequences

- Loss of ability to perform activities of daily living — dressing, eating, toileting
- Progressive dysphagia (swallowing difficulty) leading to aspiration pneumonia — the leading cause of death
- Complete loss of mobility and communication in terminal stages
- Average survival of 8–10 years after diagnosis, though ranges from 3 to 20 years

2. Brain Anatomy and Alzheimer's Disease Pathology

2.1 Key Brain Regions Affected

Understanding Alzheimer's disease requires understanding the specific brain regions it targets and why their destruction produces the characteristic symptoms. The disease does not attack the brain randomly — it follows a predictable anatomical progression, first striking memory circuits before spreading to affect the entire cortex.

Hippocampus

The hippocampus, located deep within the medial temporal lobe, is the brain's primary memory consolidation centre. It is responsible for converting short-term experiences into long-term memories and is the first major structure to be visibly damaged in Alzheimer's disease. Hippocampal volume loss of 10–15% is detectable in the pre-clinical stage, years before symptoms emerge. By the time a patient presents with memory complaints, hippocampal volume may already be reduced by 25–30%. This is why our model targets hippocampal volume as a primary MRI feature, while also recognising — critically — that hippocampal atrophy is not specific to Alzheimer's and can occur in depression, which this project explicitly accounts for.

Entorhinal cortex

The entorhinal cortex acts as the primary gateway between the hippocampus and the rest of the cortex. It is, in fact, the first region to show neurofibrillary tau tangles — even before the hippocampus. Damage here disrupts the flow of information to and from memory circuits, explaining why recent memories are lost first while remote memories are preserved.

Posterior cingulate cortex and precuneus

These regions, central to spatial memory, self-referential thinking, and episodic memory retrieval, show markedly reduced glucose metabolism (visible on FDG-PET) even in the pre-

clinical stage. They also exhibit some of the earliest and heaviest amyloid plaque deposition, making them critical targets for amyloid-PET imaging.

Prefrontal cortex

As the disease progresses, degeneration spreads to the prefrontal cortex, producing the executive dysfunction — impaired planning, reasoning, and impulse control — that characterises moderate-stage Alzheimer's.

Amygdala

The amygdala, which governs emotional memory and fear responses, is affected alongside the hippocampus in early stages. Its degeneration contributes to the emotional dysregulation, anxiety, and personality changes seen in early disease.

2.2 Neuropathological Hallmarks

Alzheimer's disease is defined at the neuropathological level by two protein aggregates that accumulate in the brain and disrupt neural function:

Amyloid-beta ($A\beta$) plaques

Amyloid-beta is a peptide fragment cleaved from amyloid precursor protein (APP). In Alzheimer's, $A\beta_{42}$ — the longer, more aggregation-prone form — misfolds and accumulates as extracellular plaques between neurons. These plaques trigger neuroinflammation and are detectable by amyloid-PET (AV45 tracer) decades before symptoms. The ratio of $A\beta_{42}$ to $A\beta_{40}$ in blood plasma is the most important blood biomarker in our dataset.

Neurofibrillary tau tangles

Tau is a microtubule-stabilising protein. In Alzheimer's, tau becomes hyperphosphorylated (p-tau), detaches from microtubules, and forms intraneuronal tangles. This destroys the neuron's internal transport system, eventually killing the cell. Tau pathology spreads in a predictable pattern (Braak staging) — starting in the entorhinal cortex and hippocampus, then spreading outward. Phosphorylated tau 181 and 217 in plasma are the most diagnostically powerful blood biomarkers.

2.3 Disease Staging — From Pre-Clinical to Terminal

Alzheimer's disease progresses through biologically distinct stages, with pathological changes beginning 15–20 years before any clinical symptoms. The table below summarises this progression:

Stage	Name	Key Changes	Biomarkers
Pre-clinical	Asymptomatic amyloid accumulation	Amyloid plaques form silently; no cognitive change; hippocampal volume normal or borderline	A β 42/40 ratio drops; p-tau181 begins rising
Stage 1 MCI	Subjective cognitive decline	Patient notices subtle changes; MMSE normal; slight hippocampal atrophy begins; others may not notice	p-tau217 elevated; amyloid PET positive
Stage 2 MCI	Mild cognitive impairment	Objective memory deficits on testing; hippocampal volume reduced 15–20%; daily function mostly intact	NfL begins rising; tau PET becomes positive in temporal lobe
Stage 3	Mild Alzheimer's dementia	Clear memory, language, and executive deficits; difficulty with complex tasks; significant cortical atrophy	All biomarkers strongly positive; FDG hypometabolism visible
Stage 4	Moderate Alzheimer's dementia	Confusion, disorientation; needs help with ADLs; behavioural changes pronounced; widespread cortical atrophy	CSF tau very elevated; extensive amyloid burden
Stage 5–6	Severe / Terminal	Loss of speech, mobility, continence; bed-bound; complete dependence; death from pneumonia or infection	Biomarkers no longer diagnostically relevant

3. Dataset — Alzheimer's Disease Neuroimaging Initiative (ADNI)

3.1 What is ADNI?

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is a longitudinal multisite study launched in 2004 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and non-profit organisations. Its primary goal is to identify neuroimaging, biological, genetic, and clinical markers of Alzheimer's disease progression, particularly the transition from normal cognition through Mild Cognitive Impairment (MCI) to AD dementia.

ADNI is the single most widely used dataset in Alzheimer's deep learning research and the de facto benchmark for any model targeting MCI-to-AD conversion prediction. Data is available to qualified researchers through the Laboratory of Neuro Imaging (LONI) image archive at adni.loni.usc.edu.

3.2 Why ADNI Was Chosen

ADNI was selected for this project for the following specific reasons, each of which directly supports the research objectives:

Reason	Relevance to this project
Multimodal data	Contains MRI, PET (amyloid and tau), CSF, blood biomarkers, genetics, and neuropsychological scores — enabling the three-modality fusion architecture
Longitudinal design	Subjects are followed at 6, 12, 24, 36, and 48-month intervals — essential for temporal trajectory modelling and MCI-to-AD conversion labelling
Diagnostic diversity	Includes cognitively normal (CN), subjective memory concern (SMC), early MCI (EMCI), late MCI (LMCI), and AD groups — covering the full disease spectrum

Scale	Over 2,000 participants across ADNI 1, GO, 2, and 3 phases — sufficient for deep learning training
Multi-site harmonisation challenge	Data from 50+ sites across North America — directly motivating the ComBat harmonisation step in preprocessing
Depression comorbidity data	Contains subjects with depression diagnoses within the CN group — enabling the explicit confound modelling between AD and depression that is a key novelty of this project
Standard benchmark	All state-of-the-art models are evaluated on ADNI — enabling direct comparison with existing literature

3.3 Dataset Composition

The following data types are used in this project:

Modality	Details
T1-weighted MRI	3D volumetric structural scans; 1mm isotropic; multiple protocols across sites; primary source of anatomical biomarkers
Amyloid PET (AV45)	Florbetapir tracer; 50–70 min post-injection frames; measures fibrillar amyloid burden
Tau PET (AV1451)	Flortaucipir tracer; measures regional tau tangle burden; available in ADNI 3
Plasma biomarkers	p-tau181, p-tau217, A β 42, A β 40, A β 42/40 ratio, NfL (neurofilament light), GFAP
APOE genotype	Number of ϵ 4 alleles (0, 1, 2) — strongest genetic risk factor for late-onset AD
Neuropsychological scores	MMSE, ADAS-Cog, CDR-SB, MoCA — used as auxiliary labels and for confound characterisation
Demographics	Age, sex, education years, family history — used as clinical covariates

4. Data Preprocessing Pipeline

Raw ADNI data requires extensive preprocessing before it can be used for model training. Each modality has a distinct processing pipeline reflecting its specific data characteristics and the biological information it encodes. The overarching principle is: preserve biologically meaningful signal while removing technical artefacts introduced by the multi-site, multi-scanner nature of ADNI.

4.1 MRI Preprocessing

Structural T1-weighted MRI preprocessing follows a standardised neuroimaging pipeline executed in the following order:

Step 1 — DICOM to NIfTI conversion

Raw ADNI MRI data is delivered in DICOM format. All downstream neuroimaging tools operate on NIfTI (.nii.gz) files. Conversion is performed using `dcm2nii`, which also extracts acquisition metadata (TR, TE, field strength, voxel size) into JSON sidecar files required for quality control.

Step 2 — N4ITK bias field correction

MRI scanners introduce a spatially smooth intensity non-uniformity — the bias field — caused by radiofrequency coil inhomogeneity. This makes the same tissue type appear brighter near the coil than far from it. N4ITK (implemented in ANTs) removes this field iteratively. This step is critical because without it, a CNN will learn coil position as a spurious feature rather than tissue anatomy. Skipping N4 correction is one of the most common preprocessing errors in published Alzheimer's MRI models.

Step 3 — Skull stripping

The skull, dura, scalp, eyes, and other non-brain tissues are removed, isolating the brain parenchyma. HD-BET (a deep learning-based tool) is preferred over the traditional FSL BET tool because it handles atrophied brains — common in MCI and AD subjects — without over-stripping temporal lobe tissue near the hippocampus. Every skull-stripped output is visually inspected, as errors in this step corrupt all downstream analyses.

Step 4 — Spatial registration to MNI152 template

Each subject's brain is registered (aligned) to the MNI152 standard space using ANTs SyNQuick, a two-stage process: first an affine (12 degrees of freedom) registration for global alignment, then a diffeomorphic non-linear SyN registration to account for individual anatomical shape differences. After registration, voxel $[x, y, z]$ corresponds to the same anatomical location across all subjects — a prerequisite for the CNN to learn spatially meaningful features.

Step 5 — Intensity normalisation

MRI intensity values are arbitrary across scans and scanners. Each brain volume is z-score normalised within its brain mask (mean subtracted, divided by standard deviation) to produce comparable intensity ranges. For multi-site ADNI data, ComBat harmonisation is additionally applied to remove site-level systematic offsets while preserving biological variance.

Step 6 — Hippocampal segmentation and ROI extraction

Subcortical structures are segmented using FastSurfer (GPU-accelerated FreeSurfer), yielding volumes for the hippocampus, entorhinal cortex, amygdala, parahippocampal gyrus, posterior cingulate, and precuneus. These ROI volumes serve dual purposes: as auxiliary tabular features fed to the clinical metadata branch, and as spatial masks for loss weighting in the imaging branches.

Step 7 — Resizing and standardisation

All volumes are resampled to 1mm isotropic voxel spacing and cropped or padded to a fixed spatial dimension of 96 x 112 x 96 voxels, covering the full MNI152 brain volume. This ensures consistent tensor dimensions across the dataset.

Step 8 — Quality control

Automated quality control using MRIQC generates image quality metrics (IQMs) including signal-to-noise ratio, contrast-to-noise ratio, and cortical-to-WM-contrast. Scans falling outside acceptable thresholds, as well as all scans with failed skull stripping or registration (assessed by visual inspection), are excluded. Historically, 5–15% of ADNI MRI scans require exclusion after QC.

4.2 PET Preprocessing

PET preprocessing differs fundamentally from MRI preprocessing because PET measures physiological activity (tracer uptake) rather than anatomy, and raw PET values are not directly comparable across subjects without normalisation.

Step 1 — DICOM to NIfTI and frame identification

PET scans are converted to NIfTI format. ADNI contains both AV45 (amyloid) and AV1451 (tau) tracers, which require different reference regions. The scan type is identified from DICOM metadata before proceeding.

Step 2 — PET-to-MRI co-registration

PET has low spatial resolution (~5-6mm FWHM) and contains no anatomical detail. Each PET scan is registered rigidly (6 degrees of freedom) to the subject's own T1-weighted MRI (already preprocessed) using ANTs. The MRI-to-MNI warp computed in MRI Step 4 is then applied to

bring the PET scan into MNI space. This two-step approach is mandatory — direct PET-to-MNI registration is unreliable due to the low PET resolution.

Step 3 — SUVR normalisation

Standardised Uptake Value ratio (SUVR) is computed by dividing each voxel's uptake by the mean uptake in a reference region known to be unaffected by Alzheimer's pathology. Reference regions differ by tracer: the whole cerebellum for AV45 amyloid PET, and the inferior cerebellar grey matter for AV1451 tau PET. SUVR normalisation makes values physiologically interpretable and comparable across subjects and scanning sessions.

Step 4 — Partial volume correction

Due to PET's limited spatial resolution, signal from small structures such as the hippocampus is contaminated by signal from surrounding tissue — the partial volume effect. The Geometric Transfer Matrix (GTM) method is used to recover true regional uptake values. This step is particularly important for medial temporal lobe structures central to Alzheimer's diagnosis.

Step 5 — Harmonisation and QC

ComBat harmonisation is applied to extracted ROI SUVR values to remove systematic scanner and site effects. Quality control confirms correct PET-MRI alignment (visually verified by overlaying both modalities), appropriate SUVR range (0.8–2.5 for amyloid PET), and correct reference region placement.

4.3 Blood Biomarker Preprocessing

Blood biomarker data from ADNI is tabular, stored in CSV files with one row per subject per visit. Preprocessing focuses on ensuring biological interpretability, handling missingness appropriately, and removing technical confounds.

Step 1 — Data loading and audit

ADNI biomarker data spans multiple CSV files across phases (ADNI1, ADNI2, ADNI3, ADNI4), each with different assay panels. All files are loaded, merged on subject ID and visit code, and audited for completeness. A subject x visit x marker presence matrix is constructed to characterise the missingness pattern before any imputation.

Step 2 — Biomarker selection

The following plasma markers are extracted based on established diagnostic validity: phosphorylated tau 181 (p-tau181), the amyloid beta 42/40 ratio (the most robust single plasma Alzheimer's biomarker), neurofilament light chain (NfL, a neurodegeneration marker), and GFAP (glial fibrillary acidic protein, an astrocyte activation marker). The Abeta42/40 ratio is always used in preference to Abeta42 alone, as it corrects for individual differences in amyloid production.

Step 3 — MICE imputation

Missing values are imputed using Multiple Imputation by Chained Equations (MICE), implemented via scikit-learn's `IterativeImputer`. MICE is chosen over simple mean imputation because it preserves the correlation structure between biomarkers — precisely the correlations (e.g., between p-tau and Abeta42/40) that carry diagnostic information and that the FT-Transformer will exploit. The imputer is fit exclusively on training data and applied to validation and test sets.

Step 4 — Z-score normalisation

Each biomarker is independently z-score normalised (mean 0, standard deviation 1) using statistics computed from the training set only. This places all biomarkers on a common scale, which is necessary for the Transformer's attention mechanism to weight them appropriately.

Step 5 — ComBat harmonisation

Plasma samples in ADNI are processed at different laboratories across multiple years. Systematic inter-lab differences (batch effects) can exceed the magnitude of the biological signal being measured. neuroCombat is applied to remove batch effects while preserving the biological covariates of interest (age, sex, diagnosis group), which are specified as protected variables.

Step 6 — Longitudinal feature engineering

For subjects with two or more visits, the rate of change (slope) of each biomarker over time is computed as an additional feature using linear regression across available timepoints. The trajectory of p-tau181, Aβ42/40, and NfL over visits is more predictive of conversion than a single cross-sectional measurement. Subjects with only one visit receive a slope of zero with a corresponding missingness indicator flag.

5. Proposed Model Architecture

The proposed system is a multimodal deep learning architecture designed to integrate structural MRI, functional PET imaging, and plasma blood biomarkers for early Alzheimer's risk prediction. The architecture is specifically engineered for clinical deployability — it produces calibrated risk scores at each modality availability level, not only when all three modalities are present.

5.1 Architectural Principles

The design follows three guiding principles that distinguish it from prior work. First, modality-specific encoding — each modality is processed by an encoder architecture matched to its data type and the features it encodes, before any cross-modality interaction occurs. Second, cross-modal attention fusion — modalities interact through learned attention mechanisms rather than simple concatenation, enabling the model to discover synergistic relationships between imaging and blood biomarkers that are diagnostically decisive. Third, missing-modality tolerance — the model is trained with random modality dropout, enabling it to function across the clinical deployment spectrum from blood-only triage to full multimodal confirmation.

5.2 Modality-Specific Encoders

MRI encoder — 3D CNN + Vision Transformer

Structural MRI is processed by a hybrid encoder. A 3D convolutional neural network (CNN) extracts local features — hippocampal volume, cortical thickness, sulcal depth — at multiple spatial scales through hierarchical pooling. The output feature maps are then tokenised and processed by a Vision Transformer (ViT) encoder, which captures long-range spatial dependencies between distant brain regions that CNN convolutions cannot reach. The CNN provides spatial precision; the Transformer provides global context. Together, they produce a set of MRI feature tokens.

PET encoder — 3D CNN + spatial self-attention

Amyloid and tau PET scans are processed by a separate 3D CNN encoder followed by spatial self-attention layers. The self-attention mechanism learns to weight amyloid-positive hotspots (posterior cingulate, precuneus, frontal cortex) more heavily than amyloid-negative regions, effectively learning to localise pathological uptake patterns without any explicit supervision on their spatial location. When both amyloid and tau PET are available, their feature tokens are concatenated before entering the fusion module.

Blood biomarker encoder — FT-Transformer

Tabular blood biomarker data is processed by a Feature Tokenisation Transformer (FT-Transformer), which first projects each scalar biomarker value into a learned embedding (tokenisation), then applies multi-head self-attention to model interactions between biomarkers. This architecture is specifically designed for tabular data and substantially outperforms MLP-based approaches on biomedical tabular datasets. Crucially, the blood encoder can disambiguate Alzheimer's-pattern atrophy from depression-pattern atrophy: p-tau181 and Abeta42/40 are elevated in Alzheimer's but not in depression, even when both conditions produce hippocampal volume loss on MRI.

5.3 Cross-Modal Attention Fusion

After independent encoding, the three modality-specific token sets are fused through a cross-modal attention module. Each modality's tokens serve as queries attending to the keys and values of the other modalities. MRI tokens query PET tokens (locating imaging correlates of amyloid burden), and both imaging token sets query blood biomarker tokens (contextualising atrophy patterns with molecular evidence). This bidirectional cross-attention enables the model to learn that hippocampal reduction combined with elevated p-tau181 and amyloid-positive PET is a very different risk signal from hippocampal reduction alone.

Key innovation

The cross-modal attention mechanism operationalises the clinical insight that motivated this project: no single biomarker is specific enough for Alzheimer's diagnosis, but their combination is. The model learns which combinations are diagnostically decisive rather than having these rules hard-coded.

5.4 Temporal Transformer

ADNI subjects are assessed at multiple visits over years. A longitudinal temporal Transformer encodes the sequence of fused multimodal representations across visits, learning the trajectory of disease progression. Position encodings encode the time interval between visits (not just visit order), allowing the model to distinguish rapid from slow progression. This temporal component is critical for early detection: the rate of biomarker change over 12–18 months is more predictive than any single cross-sectional measurement.

5.5 Output Head — Survival Analysis

Rather than producing a simple binary prediction (will convert / will not convert), the output head implements a survival analysis formulation that outputs the probability of MCI-to-AD conversion as a function of time. This produces a clinically actionable output: the probability of conversion within 12, 24, 36, and 48 months, along with a predicted median time to conversion. Survival analysis also handles the censoring problem inherent in ADNI — subjects who have not yet converted but whose follow-up data ends.

Clinical value

A clinician told 'this patient has a 78% probability of converting to Alzheimer's within 24 months' can make a fundamentally different and more useful decision than one told 'high risk'. It justifies initiating anti-amyloid therapy, enrolling in prevention trials, or implementing intensive monitoring — with a specific timeframe.

5.6 Missing Modality Tolerance

During training, entire modalities are randomly zeroed out with defined probabilities: PET with 30% probability, MRI with 20% probability, blood biomarkers with 10%. This forces the cross-modal attention mechanism to learn robust representations that do not depend on any single modality. At inference, the model accepts any available subset of modalities and produces a risk score with a confidence estimate that narrows as more modalities are provided.

This enables three deployment configurations without retraining: a blood-only triage model for GP-level screening (Stage 2 of the clinical funnel); a blood-plus-MRI confirmation model for secondary care (Stage 3); and the full three-modality model for specialist centres and clinical trials (Stage 4).

5.7 Explainability Layer

For clinical deployment, model decisions must be interpretable. The explainability layer generates three complementary outputs: GradCAM heatmaps overlaid on MRI and PET scans, highlighting the brain regions that most influenced the risk score; SHAP (SHapley Additive exPlanations) values for each blood biomarker, showing their individual contribution to the prediction; and a natural-language summary of the top contributing factors, formatted for clinical communication.

6. Research Novelty and Contributions

The following summarises the genuine novelty of this project relative to existing literature. The architecture components individually exist in the literature; the novelty lies in their integration within a clinically-grounded system design.

Contribution	Novelty and distinction from prior work
Clinically-staged missing-modality system	Existing multimodal models assume all modalities are always available. This project explicitly designs and trains for modality-incomplete inference, matching real clinical deployment where expensive PET is unavailable at triage. No existing ADNI-based paper evaluates performance systematically across modality subsets.
Explicit AD vs depression confound modelling	The hippocampal atrophy-depression confound is acknowledged in clinical literature but not modelled as an explicit training objective in deep learning. The blood biomarker branch is architected specifically to disambiguate these conditions, with depression-comorbid subjects from ADNI used as hard negative examples.
Survival analysis output for MCI conversion	Most MCI-to-AD conversion models produce binary or probabilistic single-timepoint outputs. A survival output head producing time-resolved conversion probability curves is clinically more actionable and methodologically differentiated.
Cross-modal blood-imaging attention for three modalities	Two-modality (MRI+PET) fusion is well studied. Three-modality cross-attention where blood biomarkers and imaging mutually attend to each other within a unified fusion module is rare in the Alzheimer's literature.
Biomarker trajectory encoding	Encoding the rate of change of blood biomarkers (slope across visits) as explicit features alongside raw values, within a longitudinal Transformer, extends beyond standard longitudinal models that treat visits as independent observations.

7. End-to-End Pipeline Summary

The following table summarises the complete pipeline from raw ADNI data to clinical risk prediction, showing each stage, its inputs, outputs, and the tools used.

Stage	Input	Process	Output	Tools
1. Data acquisition	ADNI portal download	DICOM to NIfTI conversion, file audit	Organised NIfTI + CSV files	dcm2niix, pandas
2. MRI preprocessing	Raw T1 MRI volumes	N4 correction, skull strip, registration, normalisation	Registered 96x112x96 volumes	ANTs, HD-BET, FSL, FastSurfer
3. PET preprocessing	Raw DICOM PET	Co-registration, SUVR, PVC, harmonisation	SUVR PET volumes in MNI space	ANTs, PETPVC, neuroCombat
4. Blood preprocessing	ADNI biomarker CSVs	MICE imputation, z-score, ComBat, slope features	Harmonised tabular biomarker array	scikit-learn, neuroCombat
5. Dataset split	All preprocessed data	GroupShuffleSplit by subject ID, stratified	Train / val / test sets	scikit-learn
6. Model training	Three modality tensors	Modality dropout, cross-modal attention, temporal Transformer	Trained model weights	PyTorch, HuggingFace
7. Inference	Available modalities (any subset)	Forward pass, survival head	Risk score + time-to-onset curve	PyTorch
8. Explainability	Trained model + patient scan	GradCAM + SHAP computation	Heatmaps + SHAP waterfall	captum, shap

8. Conclusion

Alzheimer's disease represents one of medicine's most urgent unsolved problems. Its insidious pre-clinical onset, irreversible neurodegeneration, and enormous societal burden demand detection tools that can identify patients at risk years before symptoms emerge — when intervention can still meaningfully alter the course of disease.

This project proposes a clinically-grounded multimodal deep learning system that addresses the key limitations of existing approaches: the assumption that expensive imaging is universally available, the conflation of Alzheimer's atrophy with depression-related atrophy, and the production of binary predictions that do not give clinicians actionable timelines.

By designing for the clinical reality — that most patients at-risk will first present to a GP with blood tests, that MRI and PET are reserved for those with elevated risk scores, and that what clinicians need is a time-resolved probability rather than a label — this system aims to be not merely academically novel but genuinely useful at the point of care.

The ADNI dataset provides the longitudinal, multimodal foundation required for this research. The preprocessing pipeline described in this report transforms raw, heterogeneous ADNI data into model-ready inputs while preserving the biological signal and removing technical confounds. The proposed architecture — three specialised encoders, cross-modal attention fusion, a longitudinal temporal Transformer, and a survival analysis output — operationalises the clinical insight that Alzheimer's disease can only be reliably identified through the convergence of multiple independent lines of biological evidence.

Next steps

The immediate priorities for the research team are: (1) completing the MRI and PET preprocessing pipeline on the ADNI cohort; (2) implementing and validating the FT-Transformer blood biomarker encoder; (3) training the cross-modal attention fusion module; and (4) establishing baseline performance against existing ADNI benchmarks before introducing the

novelty components.

End of Report