International Journal of Computer Science and Engineering Development
Volume 1, Issue 4 | September - October 2026 | www.ijcsed.com
ISSN: 3139-0862

MULTIMODAL AI FOR COGNITIVE SCREENING AND MRI-BASED CLASSIFICATION

Maseeha Banu[1], Hemanth S Gowda[2], Sanjay G[3], Meghana M C[4], Vandana[5]
Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning),
Malnad College of Engineering, Hassan, Karnataka, India
Corresponding author: Hemanth S Gowda, hemanthsgowdru2006@gmail.com

Abstract

This paper proposes a multimodal artificial-intelligence framework for preliminary cognitive screening and research-oriented MRI classification. The planned framework brings together fifteen cognitive games organized into five groups, speech/audio analysis, eye-movement tracking, and an MRI-based deep-learning component. The five target categories are mild cognitive impairment (MCI), Alzheimer's disease, attention-deficit/hyperactivity disorder (ADHD), Parkinson's disease, and dementia, with three proposed cognitive games associated with each category. The game groupings reflect the project's design organization and should not be interpreted as proving that any game is specific to, or diagnostic of, a particular condition. The system is intended to collect behavioural observations and, if supported by suitable datasets, examine whether MRI-derived features can contribute to classification. The integrated multimodal model, MRI pipeline, and clinical performance have not been verified as completed. Earlier project work included a prototype with a subset of the cognitive games and backend data storage; the status of each feature must be checked against the current repository before publication. This paper therefore presents the proposed architecture, planned methodology, and evaluation protocol rather than claiming experimental results. A key methodological concern is that MCI, Alzheimer's disease, and dementia are overlapping clinical concepts and cannot automatically be treated as mutually exclusive classes. Dataset labels, task definitions, and clinical interpretation must be specified before model training.

Keywords

Cognitive screening; multimodal AI; memory games; speech analysis; eye tracking; MRI; mild cognitive impairment; Alzheimer's disease; ADHD; Parkinson's disease; dementia; proposed system.

1. Introduction

Cognitive and neurological conditions can affect memory, attention, executive function, language, visual exploration, and motor performance. Digital tasks can record behavioural measures such as response time, accuracy, omissions, and error patterns. Speech and eye-movement measurements may offer additional behavioural information, while MRI provides structural brain-imaging data. A framework that evaluates these sources separately and, where valid data permit, combines them could support research into more comprehensive screening workflows.

This paper describes a proposed system built around the project's fifteen-game structure, speech/audio analysis, eye-movement tracking, and an MRI classification component. It is not presented as a clinically validated diagnostic tool. The goal is to define a technically and scientifically testable plan, clearly separating previously prototyped elements from planned development.

2. Problem Statement and Objectives

The project aims to design a modular platform that collects structured cognitive-task performance, derives selected speech and eye-movement features, and investigates MRI-based classification for five target categories. Challenges include variation in devices and task conditions, noisy behavioural measurements, limited or unmatched multimodal datasets, class imbalance, and overlap between diagnostic categories.

Objectives: (1) organize fifteen cognitive tasks into five project-defined groups; (2) record accuracy, errors, timing, and completion measures where

ISSN: 3139-0862

applicable; (3) define speech/audio feature extraction; (4) define eye-movement tracking and quality checks; (5) develop an MRI pipeline if suitable labelled data and permissions are available; (6) evaluate unimodal baselines before fusion; and (7) report only reproducible measured results after implementation.

3. Fifteen Cognitive Games and Their Project Grouping

Table 1. Project-defined grouping of the fifteen cognitive games

Target category	Game 1	Game 2	Game 3
MCI	1. Number Sequence	2. Digit Span Backward	3. 1-Back
Alzheimer's disease	4. Image Order	5. Picture Recognition	6. Object-Location
ADHD	7. Stroop	8. Go/No-Go	9. Continuous Performance Test (CPT)
Parkinson's disease	10. Trail Making	11. Tapping Speed	12. Reaction Delay
Dementia	13. Odd-One-Out	14. Category Sorting	15. Sequence Completion

The project's existing design groups fifteen tasks into five categories, with three tasks associated with each. This mapping is a design and assessment structure, not evidence that each task independently detects the corresponding disease. Performance may be influenced by age, education, language, vision, hearing, motor ability, familiarity with devices, fatigue, and other factors.

4. Proposed Multimodal Architecture

4.1 Cognitive-game module. The planned data schema can record task identifier, task version, correct/incorrect responses, completion time, attempts, omissions, and derived summary features. Task instructions and scoring rules should be fixed before evaluation. A task should not be scored using random or hard-coded values in a research experiment.

4.2 Speech/audio module. With informed consent, standardized spoken responses could be analyzed for response latency, pause duration, speaking rate, and selected acoustic or linguistic features. Automatic speech recognition and feature analysis are separate stages. Accent, language, microphone quality, background noise, and transcription error must be considered.

4.3 Eye-movement module. The proposed module may estimate gaze direction, fixation duration, and saccade-related features using a suitable camera or eye tracker. Tracking quality depends on calibration, lighting, head movement, glasses, and hardware. Ordinary camera-based tracking should not be described as medical-grade iris analysis without validation. The focus is eye-movement behaviour, not biometric identification from iris texture.

4.4 MRI module. Brain MRI data would undergo documented quality control and preprocessing, potentially including orientation checks, intensity normalization, resampling where justified, and region-of-interest or whole-volume analysis. Candidate deep-learning methods include 2D or 3D convolutional neural networks. The model and preprocessing pipeline must be selected after examining dataset size, labels, image format, acquisition sites, and permissions.

4.5 Fusion layer. Modality-specific baseline models should be evaluated first. A later late-fusion model could combine calibrated outputs from game, speech, eye-tracking, and MRI branches if compatible records exist. Missing modalities should be handled explicitly. Individual-level multimodal fusion cannot be validated by combining unrelated datasets with no participant-level linkage. Output should be described as research output, not a diagnosis.

5. Dataset and Label Strategy

The five requested categories are MCI, Alzheimer's disease, ADHD, Parkinson's disease, and dementia. These are not necessarily five mutually exclusive labels: MCI may occur in people who later develop dementia; Alzheimer's disease is a common cause of dementia; and dementia describes a syndrome that can have different causes. ADHD and Parkinson's disease also have distinct diagnostic criteria and may not be represented in the same MRI dataset as the cognitive categories.

ISSN: 3139-0862

Before implementation, select a defensible classification design. One option is to train separate dataset-specific models for conditions where suitable labels exist, then avoid presenting their outputs as directly comparable probabilities. Another is a hierarchical or multi-label design, provided the data support it. A five-class classifier should be used only if the chosen dataset has documented mutually exclusive labels suitable for that task. Do not assume one dataset contains all five conditions and all four modalities. Candidate sources to investigate include ADNI or OASIS for Alzheimer's/MCI-related research, PPMI for Parkinson's-related research, and a separately verified ADHD dataset; exact MRI availability, access terms, and labels must be checked before use.

6. Methodology and Planned Workflow

Phase 1: confirm intended use, target labels, and the fifth-condition scope (ADHD is included based on the earlier project design). Phase 2: audit datasets, licenses, diagnostic criteria, sample sizes, demographic composition, and modality availability. Phase 3: implement quality checks and standardized preprocessing for game, audio, eye-tracking, and MRI data. Phase 4: establish separate baselines for each modality. Phase 5: compare appropriate fusion methods only when matched data are available. Phase 6: conduct participant-level train/validation/test splitting to prevent the same participant's scans or recordings leaking across partitions. Phase 7: report the complete configuration, failures, and results transparently.

For game data, record the exact task implementation and scoring rules. For audio, document recording and feature extraction. For eye tracking, record calibration failures and missingness. For MRI, document preprocessing and acquisition differences. Any human-participant study requires the relevant institutional ethics review and informed consent.

7. Evaluation Plan

No new integrated-model results are claimed in this proposal. Planned classification metrics include per-class precision, recall/sensitivity, specificity where appropriate, F1-score, balanced accuracy, confusion matrices, and ROC-AUC or precision-recall AUC where suitable. Include class counts and confidence intervals where possible; overall accuracy alone is inadequate for imbalanced datasets. Evaluate probability calibration if risk scores are shown.

Compare memory-game-only, speech-only, eye-movement-only, MRI-only, and supported combined models using the same held-out participant-level test set. Report data quality, missing modality rates, tracking/calibration failures, and model failures. Where feasible, perform external validation on a separate dataset or site. Usability evaluation should assess task completion and user understanding, not claim clinical benefit without appropriate study design.

8. Technology Plan

A candidate implementation stack is Python for data processing, PyTorch or TensorFlow for deep learning, a suitable audio feature library, OpenCV or a validated eye-tracking toolkit, and the existing web/backend technologies already used in the project. The final paper must name exact packages and versions only after they are confirmed from the repository and environment. MRI preprocessing tools should be selected based on the data format and validated workflow. These are proposed choices for planned components, not evidence that all components have been built.

9. Completed Work Versus Planned Work

Earlier project discussions indicate that a prototype existed with a subset of the cognitive games and backend database storage. Historical notes also identify the fifteen-game mapping listed in Section 3. The precise current state of each game, scoring method, speech module, eye-tracking module, and backend must be verified against the present code before the manuscript claims implementation.

Planned or not verified as completed: all fifteen games working end-to-end with validated scoring; reliable speech feature extraction; calibrated eye-movement measurements; MRI dataset preparation and training; integration of all modalities; fusion model; participant-level testing; external validation; user study; and clinical evaluation. The manuscript should be updated from reproducible evidence. Do not describe a feature as completed solely because it appears in the design.

10. Expected Contributions

Expected contributions include a structured fifteen-task cognitive-game design, a modular framework

ISSN: 3139-0862

for behavioural and MRI inputs, a plan for comparing modality-specific baselines with supported fusion models, explicit treatment of overlapping diagnostic categories, and a reproducible evaluation protocol. These are intended contributions, not measured findings. Whether fusion improves performance is an empirical question.

11. Ethics, Safety, and Limitations

Cognitive-task scores, audio, eye movement, and MRI are sensitive data. Use data only under appropriate permissions, consent, and institutional rules; de-identify records; limit access; and retain only what is necessary. Language, age, education, disability, device type, and acquisition site may affect results. The system must not present a score as a confirmed diagnosis or recommend treatment. Clinical use would require further validation and appropriate professional and regulatory oversight.

The main limitation is that the complete multimodal system and MRI classifier have not been verified as implemented or tested. Matched datasets containing all modalities may be unavailable. The five categories overlap conceptually, and some conditions may require separate datasets and separate models. Therefore, no diagnostic accuracy, early-detection benefit, or clinical superiority is claimed.

12. Conclusion and Future Work

This paper proposes a multimodal research framework that combines fifteen cognitive games, speech/audio analysis, eye-movement tracking, and a planned MRI-based classification component. The project's five categories are MCI, Alzheimer's disease, ADHD, Parkinson's disease, and dementia, with three game tasks associated with each category. These groupings organize the project but do not imply that individual games diagnose a disease. Future work should verify current implementation status, choose suitable licensed datasets, define a clinically defensible label scheme, implement unimodal baselines, and only then test fusion. Results must be reported from reproducible experiments, and the platform should be described as a research prototype rather than a diagnostic system.

References — verification required before submission

1. Alzheimer's Disease Neuroimaging Initiative (ADNI), official study and data-access documentation. Verify exact release, variables, labels, and usage conditions.
2. Open Access Series of Imaging Studies (OASIS), official project and data-access documentation. Verify the specific dataset and diagnostic labels.
3. Parkinson's Progression Markers Initiative (PPMI), official study and data-access documentation. Verify MRI availability and permitted use in the selected release.
4. Identify and verify an appropriate ADHD dataset with MRI data, diagnostic labels, and clear licensing/usage terms before claiming MRI-based ADHD classification feasibility.
5. Add verified peer-reviewed literature on cognitive games, speech-based assessment, eye tracking, MRI-based classification, and multimodal learning. Check every author, title, journal, year, page range, and DOI against the original publication.

This list is a source-verification checklist, not a final bibliography. Complete the references and in-text citations before journal submission.